



AFRL-RI-RS-TR-2012-118

WIDEBAND SIGNAL DE-INTERLEAVING (WSD)

UNIVERSITY OF RHODE ISLAND

APRIL 2012

FINAL TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

STINFO COPY

**AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE**

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09. This report is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (<http://www.dtic.mil>).

AFRL-RI-RS-TR-2012-118 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/s/
PETER ZULCH
Work Unit Manager

/s/
WARREN H. DEBANY, JR., Technical Advisor
Information Exploitation & Operations Division
Information Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE*Form Approved*
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Service, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington, DC 20503.

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**1. REPORT DATE (DD-MM-YYYY)**

APR 2012

2. REPORT TYPE

Final Technical Report

3. DATES COVERED (From - To)

OCT 2008 – DEC 2011

4. TITLE AND SUBTITLE**WIDEBAND SIGNAL DE-INTERLEAVING (WSD)****5a. CONTRACT NUMBER**

FA8750-09-2-0236

5b. GRANT NUMBER

N/A

5c. PROGRAM ELEMENT NUMBER

62788F

6. AUTHOR(S)

STEVEN KAY, NARESH VANKAYALAPATI

5d. PROJECT NUMBER

459E

5e. TASK NUMBER

9B

5f. WORK UNIT NUMBER

01

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)University of Rhode Island, Research Office
70 Lower College Road
Kingston, RI 02881-1967**8. PERFORMING ORGANIZATION
REPORT NUMBER****9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**Air Force Research Laboratory/Information Directorate
Rome Research Site/RIGC
525 Brooks Road
Rome NY 13441**10. SPONSOR/MONITOR'S ACRONYM(S)**
AFRL/RI**11. SPONSORING/MONITORING
AGENCY REPORT NUMBER**
AFRL-RI-RS-TR-2012-118**12. DISTRIBUTION AVAILABILITY STATEMENT**

Approved for Public Release; Distribution Unlimited. This report is the result of contracted fundamental research deemed exempt from public affairs security and policy review in accordance with SAF/AQR memorandum dated 10 Dec 08 and AFRL/CA policy clarification memorandum dated 16 Jan 09.

13. SUPPLEMENTARY NOTES**14. ABSTRACT**

We proposed efficient techniques for passive detection and localization of emitters of low probability of intercept signals. These are the generalized likelihood ratio test (GLRT) detector and the maximum likelihood estimator (MLE). We derived the theoretical bounds for the localization variance, the Cramer-Rao lower bound (CRLB), and showed that the MLE attains the bounds. We further investigated the problems of optimal sensor configuration, knowledge aided geo-location design as well as method to improved the TDOA geo-location technique.

15. SUBJECT TERMS

Radio Frequency Geo-location, SIGINT, Detection and Geo-location

16. SECURITY CLASSIFICATION OF:**a. REPORT**
U**b. ABSTRACT**
U**c. THIS PAGE**
U**17. LIMITATION OF
ABSTRACT**

UU

**18. NUMBER
OF PAGES**

99

19a. NAME OF RESPONSIBLE PERSON
PETER ZULCH**19b. TELEPHONE NUMBER (Include area code)**
N/A

TABLE OF CONTENTS

LIST OF FIGURES	iii
LIST OF TABLES	v
ACKNOWLEDGMENTS	vi
SECTION	
1 Summary	1
2 Detection	3
2.1 Introduction	3
2.2 Methods, Assumptions, and Procedures	4
2.3 GLRT Detector	6
2.3.1 Example: A simple 2-sensor case	8
2.4 Results and Discussion	9
2.5 Some Simpler Models	13
2.5.1 Simple Bilinear Model	13
2.5.2 Classical Linear Model	14
2.5.3 Total Energy Detector	15
2.6 Conclusion	17
3 Localization	18
3.1 Introduction	18
3.2 Methods, Assumptions, and Procedures	20
3.3 CRLB and MLE of the Emitter Location	23
3.3.1 Signal unknown with unknown transmission time	23
3.3.2 Signal known with unknown transmission time	26
3.4 Results and Discussion	27
3.5 Conclusions	31
4 Optimal Sensor Configuration	32
4.1 Introduction	32
4.2 Methods, Assumptions, and Procedures	33
4.3 Results and Discussion	33
4.4 Conclusions	36
5 Knowledge Aided Design	37
5.1 Introduction	37
5.2 Methods, Assumptions, and Procedures	37
5.3 Results and Discussion	37
5.3.1 Example 1 - No Obstacles	38
5.3.2 With an Obstacle	39
5.3.3 Example 2 - A Building as an Obstacle	41
5.4 Conclusions	43
6 Improved TDOA Position Fixing	45

6.1	Introduction	45
6.2	Methods, Assumptions, and Procedures	46
6.3	Variance of the TDOAs	47
6.4	Results and Discussion	49
6.5	Conclusion	51
7	Overall Conclusion	53
	REFERENCES	55
	APPENDIX	
A	Derivation of the GLRT Detector	59
B	Localization	62
B.1	CRLB	62
B.1.1	Signal unknown with unknown transmission time	62
B.1.2	Signal known with unknown transmission time	67
B.2	Maximum Likelihood Estimator	68
B.2.1	Signal unknown with unknown transmission time	68
B.2.2	Signal known with unknown transmission time	69
B.3	Properties of $\mathbf{h}(t)$	70
B.4	Transformation of the Parameters	72
C	Optimal Sensor Configuration	75
C.1	Unconstrained Geometry	76
C.2	Constrained Geometry	77
D	Derivation of the Fisher information matrix	82
D.1	With no Obstacles	85
D.2	With an Obstacle	85
	LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS	86

LIST OF FIGURES

Figure		Page
1	Detection physical setup	10
2	Transmitted signal	11
3	Received signals	11
4	Comparison of the ROC curves	12
5	GLRT vs Maximum energy detector	13
6	Classical Linear Model vs Total Energy Detector	16
7	GLRT vs Total Energy Detector	17
8	Localization physical setup	20
9	Received Signals for Localization	22
10	Ambiguity when signal and TOA are both unknown.	24
11	LSE and the Likelihood function.	28
12	Transmitted signal waveform.	29
13	Fourier coefficients plot.	29
14	Scatter plot and 95% error ellipse	30
15	Comparison of the variances	31
16	Different Sensor Configurations.	32
17	Two D-optimal configurations for six sensors	34
18	Constrained sensor configuration setup	34
19	Constrained D-optimal configuration for three sensors	35
20	Constrained D-optimal configuration for four sensors	36
21	Physical setup and Transmitted signal	39

Figure		Page
22	Azimuth modulation function example	40
23	Setup with obstacle very close to the line of sight	41
24	Localization ellipses comparison	42
25	Setup with the obstacle away from the lines of sight	43
26	Localization ellipses comparison	44
27	Physical setup	47
28	Curvatures of the correlation function at the peak	49
29	SNR vs Mean square error	51
30	Histograms of the peak location	51
C.1	Definition of the angle ψ_i	75
C.2	Three sensor setup.	78
C.3	Four sensor setup.	80
D.1	Definition of the angle ψ_i	84

LIST OF TABLES

Table		Page
1	Comparison of computation times	12
2	CRLB for different geometries	32

ACKNOWLEDGMENTS

This work was supported by the Air Force Research Lab, Rome, NY under contract FA8750-09-2-0236. We are grateful to Dr. Peter Zulch for the continuous encouragement and the uninterrupted support. We would also like to thank Dr. Quan Ding for his contributions on the optimal sensor configuration problem. In addition we thank Dr. Tim Toolan for his help with the formatting of this report.

1 Summary

The first task of an intercept receivers is to detect the presence of a target. From the data collected at multiple platform receivers a decision has to be made on the presence or absence of a target. A common practice is to define a function of this observed data and compare the function value against a threshold to make a decision. This function, or sometimes also called as the test statistic, is called the detector. Although no explicit detector has been derived by Fowler¹, the approach taken for localization suggests that the detector would be the complex ambiguity function (CAF). We will refer to this as the CAF detector or otherwise simply as the time difference of arrival (TDOA) approach as it is commonly called. It is important to note that the detector that we are proposing is also a TDOA based detector. What we are proposing is a better technique for estimating the TDOA information.

In a nutshell the CAF detector can be explained as follows. From all the possible receiver combinations only a small subset of pairs is chosen. For the data from each of these pairs of receivers the CAF is computed. The CAF corresponding to the pair that yields the maximum CAF is taken as the test statistic and compared to a threshold. There are two drawbacks to this approach. The first is that it is using only part of the data (only a subset of all possible combinations). So the available information is not completely used which directly leads to a poor detection performance. Even if all the pairs are used there is a second drawback. The detector is depending only on the correlation factor but for detection, the energy at a particular receiver is also an important factor. This factor is not taken into account in this detector. This was also evident in some of our simulations where a plain energy detector outperformed the CAF detector.

Clearly, the currently used technique is not processing the available data efficiently. The GLRT detector that we propose processes the data more efficiently. We have shown that the GLRT detector is the maximum eigen-value of a complex ambiguity matrix. This matrix has the correlation factors from all possible combinations of receiver pairs. Also, the diagonal of this matrix is the energy at each receiver. So, we are using the two important factors - energy and correlation, in our detector and so the performance of this detector is better.

Once a target has been detected the next task of an interceptor is to localize the target. Localization is the estimation of the target position and velocity. The TDOA and frequency difference of arrival (FDOA) data can be expressed as a function of the target position and velocity. So, if the TDOA and FDOA information can be estimated from the observed data then using these estimates the target position and velocity can be calculated. In the current technique the time-difference and frequency-difference values that maximize the CAF are taken as the estimates

¹Emitter Location Processing (A Short Course) Mark L. Fowler, Dept. of Electrical and Computer Engineering, State University of New York, Binghamton, NY

for the TDOA and FDOA respectively. So for each of the chosen pairs of receivers, corresponding (TDOA, FDOA) pairs are estimated. There is a drawback to this approach. For a given target location the (TDOA,FDOA) pairs corresponding to each of the receiver pairs are all related and can take only a certain set of values. But when they are estimated independently this restriction is not applied and so the estimation is poorer. A better way to approach this is to express the CAF as a function of the target location and velocity. In our MLE we avoid these drawbacks by using all the available data and by directly estimating the target location and velocity without the intermediate step of estimating the TDOA and FDOA values.

It is often possible to use the knowledge of obstacles between the emitter and the sensors and the azimuth modulation they induce for the localization of the emitter. This is particularly useful in urban environments where there is a number of obstacles such as buildings, trees etc. These obstacles cause obstruction, reflection, diffraction etc, of the transmitted signal which induces some azimuth modulation in the signal received at the sensors. This kind of localization is called as the knowledge aided design. We have derived the theoretical results that confirm the increase in information due to these obstacles under certain scenarios. We have also analyzed the case of an obstructing obstacle.

As explained previously the TDOA approach is sub-optimal for localization. But, the advantage of the TDOA approach is that it requires very few resources. For example since only a subset of the pairs of sensors is used, data links are necessary only between those pairs. Also, the TDOAs are estimated at the local sensor pairs and this information is transmitted to the central fusion station. This requires far less bandwidth than what is required to transmit the complete signal. In situations where the resources are limited, it may not be possible to implement the MLE. Under such circumstances, the TDOA approach is the only option. We have come up with an improvement to the existing TDOA approach without requiring any additional resources. In order to estimate the TDOA at a local pair of sensors, the cross-correlation function, which is the maximum likelihood function, is maximized. We propose that the curvature of the cross-correlation function at the peak should also be estimated and then transmitted to the central fusion station along with the local TDOA estimate. The curvature of the likelihood function gives a measure of the quality of the TDOA estimate. Therefore, this curvature information can be used as a weighting factor when processing the TDOAs at the fusion station. This only requires one additional number to be transmitted to the fusion station. We have found that the curvature information is particularly useful when some of the sensors are operating at signal-to-noise ratios close to the break-down range.

In the following sections we will explain each of the above mentioned topics in detail. All the mathematical derivations are provided in the appendices.

2 Detection

2.1 Introduction

RADAR is an acronym for radio detection and ranging with detection being the crucial function. A radar system illuminates a target of interest by transmitting a signal. The echoes that are reflected back are fed to a detector which makes a decision on the presence or absence of a target. These radars are called active radars. There is another class of radars called passive radars or interceptors which silently listen for transmissions from other active devices such as active radars, jammers, beacons, etc. Intercept receivers are most desirable in hostile situations due to their covert nature [1]. In addition there is only a one-way power loss at the intercept receiver as compared to the two-way loss at the transmitting active radar. On the other hand, since the angle of signal arrival and the signal itself are generally unknown to the interceptor, efficient processing techniques such as matched filtering cannot be implemented. Also, many modern active radar systems are designed with low probability of intercept (LPI) features. They incorporate physical attributes such as frequency variability, infrequent scanning, etc. to reduce the probability of interception by an interceptor and signal design attributes such as low power, wide bandwidth, etc. to decrease the probability of detection and parameter identification at the interceptor [2]. This leads to the need for implementing highly efficient detectors in the interceptor systems.

A simple radar system with a single transmitter and a single receiver both at the same physical location is called a monostatic radar. In general, its performance is inferior to a multistatic radar system. An active multistatic radar system has one or more transmitters and many spatially separated receiving stations. A passive multistatic radar consists of only a network of distributed sensors. Such a system of multiple receive platforms has a higher probability of intercepting the signals of interest. Also, multiple platforms provide more data samples over a given interval of time which increases the probability of detection. Each of these receive platforms can perform some kind of processing on the received signal. This is called decentralized target detection [3, 4]. These processed signals from the individual receive platforms are communicated to a central processor where they are further processed to arrive at a global decision. Quite often the local receive platforms process the received signal to arrive at a local decision. These local decisions are communicated to a central processor for decision fusion. Several such decision fusion techniques have been proposed and used over the years [5, 6, 7]. Despite some practical advantages such as requiring low bandwidth data links between receive platforms and less processing power, the decentralized detection approach has an obvious performance loss due to the absence of cross-platform correlation information. On the other hand, centralized detection has better performance but requires more resources. Large bandwidth data links are required to transfer the received signals from all the receiving stations to the fusion center. High

speed signal processors are required to process the large amount of data available at the fusion center for real-time detection. Since centralized detection relies on the cross-platform correlation information, the receive platforms must be precisely synchronized in time. In [8], a centralized detector for a particular model is given. Here we use a model that is similar to the model used by Stein in [9].

When a signal is transmitted by an active device, from here on referred to as the target, it reaches each of the receiving stations after a certain amount of time, which is the propagation delay. If the target is moving with respect to the receiving station, then there is a Doppler shift in the frequency of the signal. Also, the amplitude and phase of the signal are changed due to propagation through the channel. Taking all these factors into account and assuming there is no noise, the complex envelope of the signal that is received at the i th receiving station, after sampling, is

$$\tilde{A}_i \tilde{s}[n - n_i] \exp \left(\frac{j2\pi k_i (n - n_i)}{N} \right)$$

where $\tilde{s}[n]$ is the signal emitted by the target, $\tilde{A}_i = A_i e^{j\phi_i}$ is the change in amplitude and phase, n_i is the propagation delay and k_i is the Doppler shift. (For simplicity we assume a discrete-frequency Doppler shift of $\frac{k_i}{N}$, for some large N .) Note that we are using “ \sim ” to represent complex variables. An important difference between detectors in active and passive radars is that in the active case the received signal is an echo of the transmitted signal. Thus, *the signal can be modeled as known* with unknown parameters. In the passive detection case, however, *the signal itself is unknown*. In some cases the signal is modeled as a stochastic process with unknown parameters [10]. We model the signal as deterministic and completely unknown, i.e, our signal modeling assumptions coincide with those of Stein [9].

In the next section we give a detailed description of the problem and model it as a statistical hypothesis test. In Section 2.3 we find the GLRT for the described hypothesis test. We compute the GLRT analytically for the special case of 2 sensors. Here we show that the maximum likelihood estimate (MLE) for the delay and Doppler is obtained by maximizing the CAF. In [9], Stein addressed the problem of differential delay and Doppler estimation for the case of two sensors. He also arrived at the same result that the MLE for the differential delay and Doppler is obtained by maximizing the CAF. He did not address the problem for more than two sensors. Also, he did not address the problem of target detection. In Section 2.4 we compare the performance of the GLRT against some commonly used detectors. In Section 2.5 we make further assumptions to simplify the problem model and derive the respective GLRTs. We provide conclusions in Section 2.6.

2.2 Methods, Assumptions, and Procedures

The problem we are addressing can be described as follows. We have M intercepting sensors placed at multiple locations. Each of these sensors collects N time samples in a given interval of time. The total MN samples that are collected

are available for processing at a fusion center. As is usual in practice, we assume $N > M$. We further assume that the noise at each of the receiver stations is white Gaussian and also that the noise at a receiver station is independent from noise at all the other receiver stations. For simplicity we assume that the variance of noise is $\sigma^2 = 1$. In situations where the noise does not meet these criteria, the observations can be pre-whitened [11] before processing so that the above assumptions hold. Now, the task of the detector at the fusion center is to detect the presence of an unknown signal in the MN observed samples.

Let $\tilde{s}[n], n = 0, 1, \dots, N-1$ be the n th complex time sample of the transmitted signal. A time delayed and Doppler shifted version of this signal and with a different amplitude and phase reaches the sensors. Let $\tilde{A}_i = A_i e^{j\phi_i}$ be the change in the amplitude and phase. Let n_i be the discrete time delay and k_i be the discrete Doppler shift. Let $\tilde{w}_i[n]$ be the n th complex time sample of the additive noise at the i th receiver station. If $\tilde{r}_i[n]$ is the n th time sample of the observation at the i th receiver station then we can write

$$\tilde{r}_i[n] = \tilde{A}_i \tilde{s}[n - n_i] \exp\left(\frac{j2\pi k_i(n - n_i)}{N}\right) + \tilde{w}_i[n] \quad n = 0, 1, \dots, N-1 \\ i = 0, 1, \dots, M-1.$$

When there is no signal, the observation is just noise, $\tilde{r}_i[n] = \tilde{w}_i[n]$. So, the hypothesis test for the detection problem can be written as

$$\mathcal{H}_0 : \tilde{r}_i[n] = \tilde{w}_i[n] \\ \mathcal{H}_1 : \tilde{r}_i[n] = \tilde{A}_i \tilde{s}[n - n_i] \exp\left(\frac{j2\pi k_i(n - n_i)}{N}\right) + \tilde{w}_i[n] \quad n = 0, 1, \dots, N-1 \\ i = 0, 1, \dots, M-1.$$

Notice that here we are modeling the received signal as unknown but deterministic with unknown parameters $\tilde{A}_i, n_i, k_i, i = 0, 1, \dots, M-1$ and $\tilde{s}[n], n = 0, 1, \dots, N-1$. Now, if we let $\omega = \exp(\frac{j2\pi}{N})$ and $\tilde{\mathbf{W}}$ be the $N \times N$ matrix $\tilde{\mathbf{W}} = \text{diag}(\omega^0, \omega^1, \dots, \omega^{N-1})$ and let \mathbf{P} be an $N \times N$ permutation matrix defined as $[\mathbf{P}]_{ij} = 1$ if $i = j + 1$ and 0 otherwise, $i = 0, 1, \dots, N-1, j = 0, 1, \dots, N-1$ and $[\mathbf{P}]_{0,N-1} = 1$, then we can write the hypothesis test in vector form as follows

$$\mathcal{H}_0 : \tilde{\mathbf{s}} = \mathbf{0} \\ \mathcal{H}_1 : \tilde{\mathbf{s}} \neq \mathbf{0}$$

where

$$\tilde{\mathbf{r}}_i = \tilde{A}_i \mathbf{P}^{n_i} \tilde{\mathbf{W}}^{k_i} \tilde{\mathbf{s}} + \tilde{\mathbf{w}}_i \quad i = 0, 1, \dots, M-1$$

and $\tilde{\mathbf{r}}_i = [\tilde{r}_i[0] \ \tilde{r}_i[1] \ \dots \ \tilde{r}_i[N-1]]^T$, $\tilde{\mathbf{s}} = [\tilde{s}[0] \ \tilde{s}[1] \ \dots \ \tilde{s}[N-1]]^T$ and $\tilde{\mathbf{w}}_i = [\tilde{w}_i[0] \ \tilde{w}_i[1] \ \dots \ \tilde{w}_i[N-1]]^T$. The permutation matrix \mathbf{P} circularly shifts $\tilde{\mathbf{s}}$. This causes an effect as if $\tilde{\mathbf{s}}$ was periodic with period equal to N samples. So, as in [9], we assume that the signal $\tilde{\mathbf{s}}$ is non-zero only in an interval that is smaller than N samples and that the discrete-time delays are relatively small compared to N . Now let $\tilde{\mathbf{r}} = [\tilde{\mathbf{r}}_0^T \ \tilde{\mathbf{r}}_1^T \ \dots \ \tilde{\mathbf{r}}_{M-1}^T]^T$, $\tilde{\mathbf{w}} = [\tilde{\mathbf{w}}_0^T \ \tilde{\mathbf{w}}_1^T \ \dots \ \tilde{\mathbf{w}}_{M-1}^T]^T$,

$$\tilde{\mathbf{A}} = \begin{bmatrix} \tilde{A}_0 & \tilde{A}_1 & \cdots & \tilde{A}_{M-1} \end{bmatrix}^T, \mathbf{n} = \begin{bmatrix} n_0 & n_1 & \cdots & n_{M-1} \end{bmatrix}^T, \\ \mathbf{k} = \begin{bmatrix} k_0 & k_1 & \cdots & k_{M-1} \end{bmatrix}^T \text{ and}$$

$$\tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) = \begin{bmatrix} (\tilde{A}_0 \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0})^T & (\tilde{A}_1 \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1})^T & \cdots & (\tilde{A}_{M-1} \mathbf{P}^{n_{M-1}} \tilde{\mathbf{W}}^{k_{M-1}})^T \end{bmatrix}^T.$$

The hypothesis test can be written as follows

$$\begin{aligned} \mathcal{H}_0 : \tilde{\mathbf{s}} &= \mathbf{0} \\ \mathcal{H}_1 : \tilde{\mathbf{s}} &\neq \mathbf{0} \end{aligned}$$

where

$$\tilde{\mathbf{r}} = \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k})\tilde{\mathbf{s}} + \tilde{\mathbf{w}}$$

and $\tilde{\mathbf{w}}$ has a complex normal distribution with zero mean and the identity matrix as the covariance matrix, i.e, $\tilde{\mathbf{w}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{MN})$, \mathbf{I}_{MN} is an $MN \times MN$ identity matrix. Here $\tilde{\mathbf{A}}$ is $M \times 1$, $\tilde{\mathbf{s}}$ is $N \times 1$, \mathbf{n} is $M \times 1$, \mathbf{k} is $M \times 1$, and are all assumed unknown.

2.3 GLRT Detector

In hypothesis testing problems where the probability density functions (PDFs) under both the hypotheses are completely known the Neyman-Pearson (NP) detector is the uniformly most powerful (UMP) detector [12]. When there are unknown parameters in the PDFs, the performance of the NP detector depends on the true value of these parameters. There are two common ways to deal with these unknown parameters. They can be modeled as random variables with some PDF and then integrated out or they can be modeled as unknown but deterministic and replaced with their MLE. In the GLRT the unknown parameters are replaced by their MLEs under the different hypotheses. Asymptotically, the GLRT is the UMP test among all tests that are invariant [12]. The likelihood ratio for the previously described hypotheses test is given by

$$L(\tilde{\mathbf{r}}) = \frac{p(\tilde{\mathbf{r}}; \tilde{\mathbf{s}}, \tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}, \mathcal{H}_1)}{p(\tilde{\mathbf{r}}; \mathcal{H}_0)} \quad (1)$$

If we replace the unknown parameters $\tilde{\mathbf{A}}$, $\tilde{\mathbf{s}}$, \mathbf{n} and \mathbf{k} with their respective MLEs $\hat{\tilde{\mathbf{A}}}$, $\hat{\tilde{\mathbf{s}}}$, $\hat{\mathbf{n}}$ and $\hat{\mathbf{k}}$ then the GLRT decides \mathcal{H}_1 if

$$L_G(\tilde{\mathbf{r}}) = \frac{p(\tilde{\mathbf{r}}; \hat{\tilde{\mathbf{s}}}, \hat{\tilde{\mathbf{A}}}, \hat{\mathbf{n}}, \hat{\mathbf{k}}, \mathcal{H}_1)}{p(\tilde{\mathbf{r}}; \mathcal{H}_0)} > \gamma \quad (2)$$

As derived in Appendix A, the GLRT test statistic is

$$\ln L_G(\tilde{\mathbf{r}}) = \max_{\mathbf{n}, \mathbf{k}} \lambda_{\max}(\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k})) \quad (3)$$

where λ_{\max} is the maximum eigenvalue and $\tilde{\mathbf{B}}$ is the $M \times M$ complex cross-ambiguity matrix (CAM) given by

$$[\tilde{\mathbf{B}}]_{ij} = \tilde{\mathbf{r}}_i^H \mathbf{P}^{n_i} \tilde{\mathbf{W}}^{k_i} (\tilde{\mathbf{W}}^{k_j})^H (\mathbf{P}^{n_j})^H \tilde{\mathbf{r}}_j \quad (4)$$

where H is conjugate transpose. Since $\tilde{\mathbf{B}}$ is positive definite, the maximum eigenvalue is real and positive. Therefore the GLRT decides that a signal is present if the maximum eigenvalue of the CAM, when also maximized over time delay and Doppler, is greater than a threshold. It should be noticed that the diagonal elements in the CAM are the energy terms at each of the receivers, i.e., $\tilde{\mathbf{r}}_i^H \tilde{\mathbf{r}}_i$, and the off-diagonal terms are the CAF values between all the pairs of sensors. For two finite length discrete time complex signals $\tilde{r}_0[n]$ and $\tilde{r}_1[n]$, $n = 0, 1, \dots, N-1$ the energy in each of the signals is given by $\mathcal{E}_i = \sum_{n=0}^{N-1} |\tilde{r}_i[n]|^2$, $i = 0, 1$ respectively and the CAF is a two dimensional function of differential delay (Δn) and differential Doppler shift (Δk) between the signals and is defined as

$$CAF(\Delta n, \Delta k) = \sum_{n=0}^{N-1} \tilde{r}_0[n] \tilde{r}_1^*[n + \Delta n] \exp\left(\frac{j2\pi\Delta kn}{N}\right) \quad (5)$$

where $*$ represents complex conjugate. It is important to note that the CAM contains CAF terms for all possible sensor pair combinations and not just a selected set of pairs. So, the CAM can also be written as $[\tilde{\mathbf{B}}]_{ij} = \mathcal{E}_i$ if $i = j$ and $[\tilde{\mathbf{B}}]_{ij} = CAF_{ij}$ if $i \neq j$ where CAF_{ij} is the CAF of the observations at sensor i and sensor j . Another interesting result is that when the correlation information is zero, all the off diagonal terms become zero and so the maximum eigenvalue is simply the maximum of the energies of the sensors. Hence the GLRT simply reduces to a type of energy detector.

At first glance, the maximization of the eigenvalue appears to be on a $2M$ dimensional space. This computation can be prohibitive when the number of sensors is large. But it should be noted that the true values of the delay and Doppler lie in a much smaller space. This is explained by the relation of the delay and Doppler to the target location and velocity. If we assume that the location and velocity of the sensors is known, then the delays to the sensors are a function of the target location and the Doppler shifts are a function of the target location and velocity. So, if (x_T, y_T, z_T) are the three dimensional coordinates of the target location and (v_x, v_y, v_z) are the target velocity components in the x, y and z directions respectively, then we can write the delays as $\mathbf{n}(x_T, y_T, z_T)$ and the Doppler shifts as $\mathbf{k}(x_T, y_T, z_T, v_x, v_y, v_z)$. Putting these back in equation (3) we have

$$\ln L_G(\tilde{\mathbf{r}}) = \max_{x_T, y_T, z_T, v_x, v_y, v_z} \lambda_{\max}(\tilde{\mathbf{B}}(x_T, y_T, z_T, v_x, v_y, v_z)) \quad (6)$$

Hence, the maximization is at most on a six dimensional space and can be performed numerically. Any additional information about target location and velocity, viz. target is on the ground or target is stationary, further reduces this space. Notice that the $(\hat{r}_T, \hat{y}_T, \hat{z}_T, \hat{v}_x, \hat{v}_y, \hat{v}_z)$ that maximize the test statistic are *the MLEs of*

the target position and velocity. Therefore, we have simultaneously estimated the target location and velocity as well. We will investigate the localization problem in section 3. Weiss investigated the problem of localization of narrowband radio frequency transmitters in [13] and arrived at a similar result.

The fact that the GLRT test statistic is a function of the maximum eigenvalue of the CAM can be further understood as follows. The problem can be viewed as a rank one approximation of an observation matrix. Let $\tilde{\mathbf{R}}$ be the $N \times M$ observation matrix given by $\tilde{\mathbf{R}} = [\tilde{\mathbf{r}}_0 \quad \tilde{\mathbf{r}}_1 \quad \cdots \quad \tilde{\mathbf{r}}_{M-1}]$. Assuming there are zero delays and Doppler shifts, we have

$$\tilde{\mathbf{R}} = \tilde{\mathbf{s}}\tilde{\mathbf{A}}^T + \tilde{\mathbf{W}}_N \quad (7)$$

where $\tilde{\mathbf{W}}_N$ is the $N \times M$ noise matrix given by $\tilde{\mathbf{W}}_N = [\tilde{\mathbf{w}}_0 \quad \tilde{\mathbf{w}}_1 \quad \cdots \quad \tilde{\mathbf{w}}_{M-1}]$. Note that the signal component $\tilde{\mathbf{s}}\tilde{\mathbf{A}}^T$ is a rank one matrix with one singular value $= \sqrt{\tilde{\mathbf{s}}^H \tilde{\mathbf{s}} \tilde{\mathbf{A}}^H \tilde{\mathbf{A}}} = \sqrt{\lambda_{\max}}$ so that the maximum eigenvalue is $\lambda_{\max} = \tilde{\mathbf{s}}^H \tilde{\mathbf{s}} \tilde{\mathbf{A}}^H \tilde{\mathbf{A}}$. The noise $\tilde{\mathbf{W}}_N$ causes $\tilde{\mathbf{R}}$ to be full rank. Hence, the detector attempts to extract a characteristic of the signal. It is conjectured that in the presence of multiple targets, say n targets, the test statistics will be the maximum n eigenvalues of the CAM.

2.3.1 Example: A simple 2-sensor case

We have seen that the GLRT statistic is a function of the maximum eigenvalue. In general, it is difficult to compute the eigenvalues analytically, but for the special case of 2-sensors it is possible. When $M = 2$ the CAM is a 2×2 matrix and so the eigenvalues are computed as follows. We have

$$\begin{aligned} \tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k}) &= \begin{bmatrix} \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} \\ \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} \end{bmatrix} \cdot \begin{bmatrix} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 & (\tilde{\mathbf{W}}^{k_1})^H (\mathbf{P}^{n_1})^H \tilde{\mathbf{r}}_1 \end{bmatrix} \\ &= \begin{bmatrix} \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 & \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} (\tilde{\mathbf{W}}^{k_1})^H (\mathbf{P}^{n_1})^H \tilde{\mathbf{r}}_1 \\ \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 & \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_1})^H (\mathbf{P}^{n_1})^H \tilde{\mathbf{r}}_1 \end{bmatrix} \end{aligned}$$

Since $\mathbf{P}^{n_i} (\mathbf{P}^{n_i})^H = \mathbf{I}_N$ (note that $\mathbf{P}^T \mathbf{P} = \mathbf{I}_N$) and $\tilde{\mathbf{W}}^{k_i} (\tilde{\mathbf{W}}^{k_i})^H = \mathbf{I}_N$, where \mathbf{I}_N is an $N \times N$ identity matrix, we have

$$\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k}) = \begin{bmatrix} \tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_0 & \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} (\tilde{\mathbf{W}}^{k_1})^H (\mathbf{P}^{n_1})^H \tilde{\mathbf{r}}_1 \\ \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 & \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_1 \end{bmatrix}$$

Notice that the diagonal elements are the energies at individual sensors and the off diagonal terms are the cross ambiguity terms. Since $\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k})$ is a 2×2 matrix it has two eigenvalues which can be determined analytically. The maximum of the

two real and positive eigenvalues is

$$\lambda_{\max}(\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k})) = \frac{\tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_0 + \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_1 + \sqrt{[\tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_0 - \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_1]^2 + 4 \left| \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 \right|^2}}{2} \quad (8)$$

In order to maximize $\lambda_{\max}(\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k}))$ over (\mathbf{n}, \mathbf{k}) , we have to maximize

$$\left| \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 \right|.$$

If we let $\Delta k = k_1 - k_0$ and $\Delta n = n_1 - n_0$ then, since $\mathbf{P}^{n_i} (\mathbf{P}^{n_j})^H = \mathbf{P}^{(n_i - n_j)}$ and $\tilde{\mathbf{W}}^{k_i} (\tilde{\mathbf{W}}^{k_j})^H = \tilde{\mathbf{W}}^{(k_i - k_j)}$, we have

$$\begin{aligned} \max_{\mathbf{n}, \mathbf{k}} \left| \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 \right| &= \max_{\Delta n, \Delta k} \left| \tilde{\mathbf{r}}_1^H \mathbf{P}^{\Delta n} \tilde{\mathbf{W}}^{\Delta k} \tilde{\mathbf{r}}_0 \right| \\ &= \max_{\Delta n, \Delta k} \left| \sum_{n=0}^{N-1} \tilde{r}_0[n] \tilde{r}_1^*[n + \Delta n] \exp\left(\frac{j2\pi \Delta k n}{N}\right) \right| \end{aligned} \quad (9)$$

The above expression is the magnitude of the CAF in discrete time and the $\Delta \hat{n}$ and $\Delta \hat{k}$ that maximize it are the MLEs of Δn and Δk respectively. A similar result was derived in [9] for the 2 sensor case in continuous time, but the CAF was maximized only to obtain the MLEs of Δn and Δk and the detection problem was not addressed. In [14] however, Holt derived a similar expression for the GLRT. He showed that when the signal is modeled as deterministic and completely unknown, the GLRT test statistic is a weighted sum of the energies in the observations at each sensor plus the real part of the appropriately weighted cross ambiguity function. Here we have derived the GLRT test statistic to be, from (8) and (9)

$$\ln L_G(\tilde{\mathbf{r}}) = \frac{\tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_0 + \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_1}{2} + \sqrt{\left(\frac{\tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_0 - \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_1}{2} \right)^2 + \max_{\Delta n, \Delta k} |CAF(\Delta n, \Delta k)|^2} \quad (10)$$

where $CAF(\Delta n, \Delta k) = \sum_{n=0}^{N-1} \tilde{r}_1^*[n + \Delta n] \tilde{r}_0[n] \exp\left(\frac{j2\pi \Delta k n}{N}\right)$ and is the CAF previously given in (5). Therefore, the GLRT incorporates the CAF as well as the energy information.

2.4 Results and Discussion

We compared the performance of the GLRT against two commonly used detectors. One detector computes the energy of the observations individually at each sensor and when the energy at any one sensor exceeds a predetermined threshold, the target is declared as present. In mathematical terms, if $\mathcal{E}_i, i = 0, 1, \dots, M-1$ are the energies of the observations at the M sensors, then the detector decides that a target is present if $\max\{\mathcal{E}_0, \mathcal{E}_1, \dots, \mathcal{E}_{M-1}\} > \gamma_\varepsilon$. This is called the maximum energy detector. It can be noticed that this detector is based solely on the energy

information. The other detector is one that is based solely on the cross-sensor correlation information. Here a sensor is fixed as the reference sensor and the CAFs are computed between observations from the reference sensor and all the other sensors. When the maximum magnitude over possible delay and Doppler shifts of any one of these CAFs exceeds a threshold, the target is declared as present. In mathematical terms, assume sensor 0 is the reference sensor and $\Delta n_i = (n_i - n_0)$ and $\Delta k_i = (k_i - k_0)$ are the difference in delay and Doppler at sensor i and sensor 0 respectively. Then if

$$|CAF_i| = \max_{\Delta n_i, \Delta k_i} \left| \sum_{n=0}^{N-1} \tilde{r}_0[n] \tilde{r}_i^*[n + \Delta n_i] \exp \left(\frac{j2\pi \Delta k_i n}{N} \right) \right|, \quad i = 1, 2, \dots, M-1$$

are the maximum magnitudes over delay and Doppler of the CAFs between sensor 0 and all the other sensors, then a target is declared as present if

$$\max\{|CAF_1|, |CAF_2|, \dots, |CAF_{M-1}|\} > \gamma_{CAF}.$$

This is referred to as pair-wise maximum CAF detector.

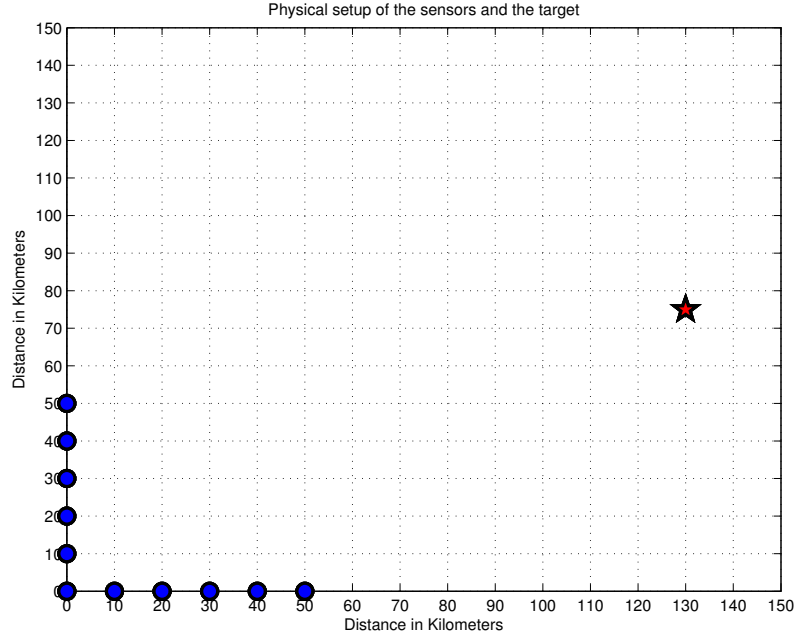


Figure 1. Physical placement of the sensors and the target position used for simulation.

For the purpose of simulation a set of 11 sensors were placed in a configuration as shown in Figure 27. To simplify the computations we assumed that the target is stationary and so the Doppler shift is zero. However, delays are incorporated. We used a Gaussian pulse for the signal which is shown in Figure 12. The length of the signal in time is $10 \mu s$ and its bandwidth is approximately 0.5 MHz. The

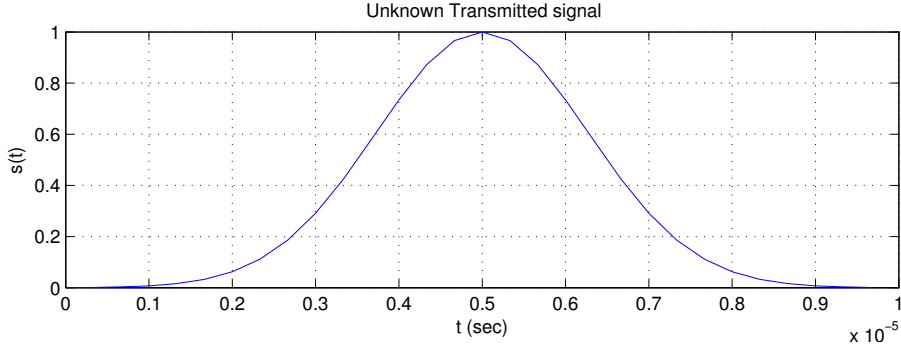


Figure 2. A Gaussian pulse that is used as the unknown transmitted signal

signal was sampled at a rate of $F_s = 3$ MHz to collect 30 non-zero samples. A total observation interval of $N = 450$ samples was required in order to allow the signal to reach all the sensors. This is shown in Figure 9. White Gaussian noise

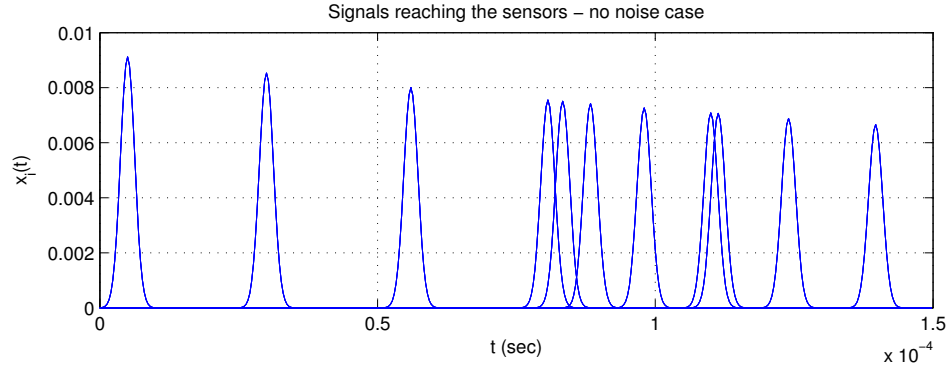


Figure 3. Signal reaching different sensors at different times with different attenuations.

was used as the additive noise at the sensors. The noise variance was adjusted so that the average energy-to-noise ratio (AENR) is 10 dB. The AENR is the ratio of the energy in the signal to the noise power at each sensor, averaged over all the sensors, i.e., if $\mathcal{E}_{si} = |\tilde{A}_i|^2 \sum_{n=0}^{N-1} |\tilde{s}[n]|^2$ is the energy of the signal at the i th sensor and σ_i^2 is the noise variance at the i th sensor, then the signal to noise ratio averaged over M sensors is given by $10 \log \left(\frac{1}{M} \sum_{i=0}^{M-1} \frac{\mathcal{E}_{si}}{\sigma_i^2} \right)$. Here we assumed a single target is located at (130,75) km and the maximization of the GLRT statistic was done over (x, y) . We used a grid search to maximize the test statistic. The comparison receiver operating characteristics (ROC) curves are shown in Figure 4. We ran 1000 simulations for each of the detectors to generate the ROC curves. The grid search was performed over a grid of size $3 \text{ km} \times 3 \text{ km}$ around the true target location with a grid-point distance of 62.5 m. The computation times in MATLAB for the pair-wise maximum CAF detector, maximum energy detector and the GLRT are given in Table 2.4.

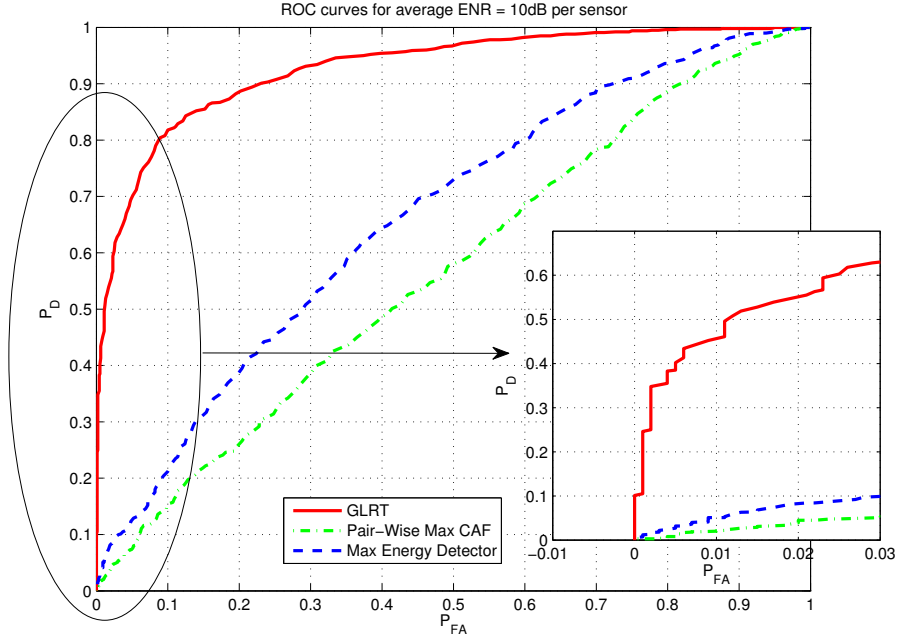


Figure 4. Comparison of GLRT against the maximum energy detector and the pair-wise maximum CAF detector

Table 1. Comparison of the computation times for the three detectors.

pair-wise maximum CAF detector	maximum energy detector	GLRT
7.18 sec	1.58 sec	778.78 sec

Incorporation of the Doppler will increase the computation time further and hence has not been included at this time. Usage of efficient maximization techniques may reduce the computation time at which point the Doppler parameter could be included. At this point we have investigated the GLRT itself and have left the details of its efficient numerical computation for a future paper. It can be noticed that the performance of the GLRT is very much better than either of the two commonly used detectors. For a probability of false alarm $P_{FA} = 0.01$ the probability of detection P_D is about 0.05 for the maximum energy detector and 0.02 for the pair-wise maximum CAF detector but is 0.45 for the GLRT. Therefore, the GLRT which is a function of the energy at each sensor combined with the cross-sensor correlation is better than detectors that are solely based on either energy or cross-sensor correlation. Altes arrives at the same conclusion in [15]. His results indicate that cross-correlation alone is usually sub optimum. He concludes that detection should use a weighted sum of pairwise cross-correlation between subarrays and energy detection at each subarray.

Next, in Figure 5, the same emitter signal was used for the case of $M = 2$

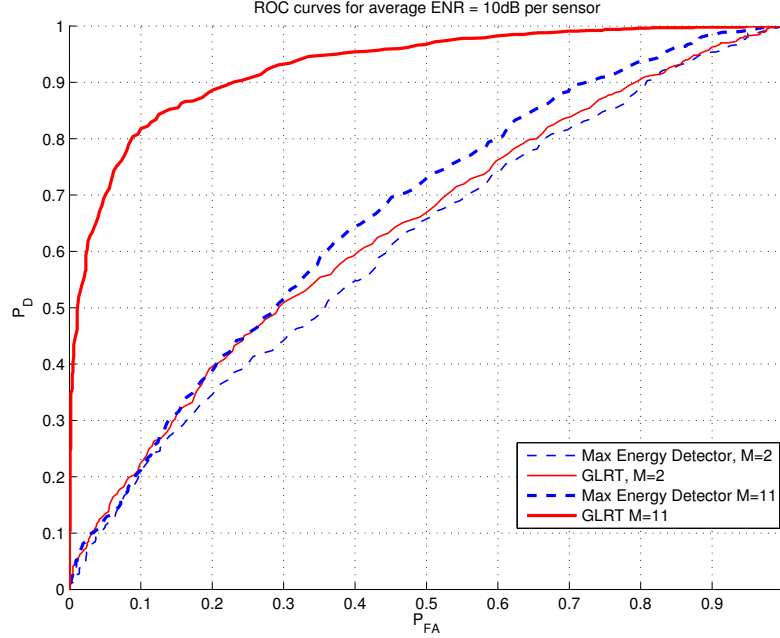


Figure 5. GLRT vs maximum energy detector for the two cases of $M=2$ and $M=11$.

and compared against the case of $M = 11$. It should be noticed that when $M = 2$ the GLRT does only slightly better than the energy detector as there is very little cross-correlation information but when M is increased to 11 the performance of GLRT exceeds that of energy detector considerably. The AENR at each sensor was set to 10 dB. In real world, as the signal travels from the target to the sensor the signal power is attenuated, which is called the propagation loss. In our simulation, while generating the observations at each sensor, we accounted for this propagation loss also when computing the \tilde{A}_i s by making the \tilde{A}_i s inversely proportional to the distance.

2.5 Some Simpler Models

The GLRT detector that was previously derived is a complete solution accounting for time delays and Doppler shifts. Depending on the specific problem appropriate assumptions can be made in the GLRT to arrive at simpler models that can be more easily implemented.

2.5.1 Simple Bilinear Model

Assume a situation where the target and the sensors are stationary - hence there is no Doppler. Also, assume that if a target is present, its location is a priori known. In such cases the time delay and Doppler parameters can be dropped and a simple bilinear (multiplicative) model [16] can be used. For such a model the

signal model and hypothesis test are given by

$$\begin{aligned}\tilde{\mathbf{r}} &= \tilde{\mathbf{H}}(\tilde{\mathbf{A}})\tilde{\mathbf{s}} + \tilde{\mathbf{w}} \\ \mathcal{H}_0 : \tilde{\mathbf{s}} &= \mathbf{0} \\ \mathcal{H}_1 : \tilde{\mathbf{s}} &\neq \mathbf{0}\end{aligned}$$

where $\tilde{\mathbf{H}}(\tilde{\mathbf{A}}) = [\tilde{A}_0\mathbf{I}_N \quad \tilde{A}_1\mathbf{I}_N \quad \cdots \quad \tilde{A}_{M-1}\mathbf{I}_N]$. Dropping the delay and Doppler terms in equation (3), we have the GLRT test statistic for this model as

$$\ln L_G(\tilde{\mathbf{r}}) = \lambda_{\max}(\tilde{\mathbf{B}})$$

where λ_{\max} is the maximum eigenvalue and $\tilde{\mathbf{B}}$ is cross-sensor correlation matrix given by

$$\tilde{\mathbf{B}} = \begin{bmatrix} \tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_0 & \tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_1 & \cdots & \tilde{\mathbf{r}}_0^H \tilde{\mathbf{r}}_{M-1} \\ \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_0 & \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_1 & \cdots & \tilde{\mathbf{r}}_1^H \tilde{\mathbf{r}}_{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{\mathbf{r}}_{M-1}^H \tilde{\mathbf{r}}_0 & \tilde{\mathbf{r}}_{M-1}^H \tilde{\mathbf{r}}_1 & \cdots & \tilde{\mathbf{r}}_{M-1}^H \tilde{\mathbf{r}}_{M-1} \end{bmatrix}$$

Notice that the principal diagonal elements of $\tilde{\mathbf{B}}$ are the energies at each of the sensors. The off-diagonal elements are the cross-correlation terms. These results are analogous to the results in Section 2.3 except that here $\tilde{\mathbf{B}}$ is not a function of the delay and Doppler.

2.5.2 Classical Linear Model

This assumes furthermore that the signal arriving at the sensors has the same amplitude and phase which we incorporate into the unknown signal $\tilde{\mathbf{s}}$. So, assuming $\tilde{A}_i = 1$ for $i = 0, 1, \dots, M-1$, we have the classical linear model as

$$\tilde{\mathbf{r}} = \mathbf{H}\tilde{\mathbf{s}} + \tilde{\mathbf{w}}$$

where the $MN \times N$ matrix $\mathbf{H} = [\mathbf{I}_N \quad \mathbf{I}_N \quad \cdots \quad \mathbf{I}_N]^T$ is known and $\tilde{\mathbf{r}}, \tilde{\mathbf{s}}, \tilde{\mathbf{w}}$ are same as defined in Section 2.2. The hypothesis test for this model is given by

$$\begin{aligned}\mathcal{H}_0 : \tilde{\mathbf{s}} &= \mathbf{0} \\ \mathcal{H}_1 : \tilde{\mathbf{s}} &\neq \mathbf{0}\end{aligned}$$

The GLRT for this hypothesis is to decide \mathcal{H}_1 if [11]

$$T(\tilde{\mathbf{r}}) = \frac{\hat{\mathbf{s}}^H (\mathbf{H}^H \mathbf{H}) \hat{\mathbf{s}}}{1/2} > \gamma'$$

where

$$\hat{\mathbf{s}} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H \tilde{\mathbf{r}} = \frac{1}{M} \sum_{i=0}^{M-1} \tilde{\mathbf{r}}_i \quad (11)$$

is the MLE of $\tilde{\mathbf{s}}$ under \mathcal{H}_1 (note that $\mathbf{H}^H \mathbf{H} = M \mathbf{I}_N$). Combining the two equations, we have

$$T(\tilde{\mathbf{r}}) = \frac{\sum_{i=0}^{M-1} \hat{\mathbf{s}}^H \tilde{\mathbf{r}}_i}{1/2} = \frac{\sum_{i=0}^{M-1} \left(\sum_{n=0}^{N-1} \tilde{r}_i[n] \hat{\mathbf{s}}^*[n] \right)}{1/2} \quad (12)$$

which is an estimator-correlator summed over all the sensors. When the additive noise is Gaussian, the $T(\tilde{\mathbf{r}})$ has a central chi-squared distribution under \mathcal{H}_0 and a noncentral chi-squared distribution under \mathcal{H}_1 . The exact detection performance is given by

$$\begin{aligned} P_{FA} &= Q_{\chi_{2N}^2}(\gamma') \\ P_D &= Q_{\chi_{2N}^2(\lambda)}(\gamma') \end{aligned} \quad (13)$$

where $Q_{\chi_{2N}^2}$ is the right-tail probability of a random variable with central chi-squared distribution with $2N$ degrees of freedom and $Q_{\chi_{2N}^2(\lambda)}$ is the right-tail probability of a random variable with noncentral chi-squared distribution with $2N$ degrees of freedom and with λ as the noncentrality parameter. The noncentrality parameter is

$$\lambda = \tilde{\mathbf{s}}^H (\mathbf{H}^H \mathbf{H}) \tilde{\mathbf{s}} = M \tilde{\mathbf{s}}^H \tilde{\mathbf{s}} \quad (14)$$

which is the total signal energy at all the sensors.

2.5.3 Total Energy Detector

In Section 2.4 we have seen the maximum energy detector that is commonly used where the energies at each of the sensors are independently compared against respective thresholds. This is decentralized detection. If the energies computed at each sensor are all summed at a fusion center and used as a test statistic, then we have the *total* energy detector. The fact that the signal received at each of the sensors is originating from the same source is ignored here also in the detector design. If $\tilde{\mathbf{s}}_i = [\tilde{s}_i[0] \ \tilde{s}_i[1] \ \cdots \ \tilde{s}_i[N-1]]^T$ is the source signal at the i th sensor and $\tilde{\mathbf{s}} = [\tilde{\mathbf{s}}_0^T \ \tilde{\mathbf{s}}_1^T \ \cdots \ \tilde{\mathbf{s}}_{M-1}^T]^T$, then using the same definitions for $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{w}}$ as in Section 2.2, we have a classical linear model for this problem as

$$\tilde{\mathbf{r}} = \tilde{\mathbf{s}} + \tilde{\mathbf{w}}$$

Notice that here, unlike the previous case, we have modeled the source signal for each sensor as a different unknown parameter. The hypothesis test for this model can be written as

$$\begin{aligned} \mathcal{H}_0 &: \tilde{\mathbf{s}} = \mathbf{0} \\ \mathcal{H}_1 &: \tilde{\mathbf{s}} \neq \mathbf{0} \end{aligned}$$

The GLRT for this hypothesis test is to decide \mathcal{H}_1 if [11]

$$T(\tilde{\mathbf{r}}) = \frac{\hat{\mathbf{s}}^H \hat{\mathbf{s}}}{1/2} > \gamma''$$

where $\hat{\tilde{\mathbf{s}}} = \tilde{\mathbf{r}}$ is the MLE of $\tilde{\mathbf{s}}$ under \mathcal{H}_1 . Putting this in the above equation, we have

$$T(\tilde{\mathbf{r}}) = \frac{\tilde{\mathbf{r}}^H \tilde{\mathbf{r}}}{1/2} = \frac{\sum_{i=0}^{M-1} \sum_{n=0}^{N-1} |\tilde{r}_i[n]|^2}{1/2} \quad (15)$$

and should be compared to (12). When the additive noise is Gaussian, here also the $T(\tilde{\mathbf{r}})$ has a central chi-squared distribution under \mathcal{H}_0 and a noncentral chi-squared distribution under \mathcal{H}_1 . But the degrees of freedom here is $2MN$. The exact detection performance of this detector is given by

$$\begin{aligned} P_{FA} &= Q_{\chi^2_{2MN}}(\gamma'') \\ P_D &= Q_{\chi'^2_{2MN}(\lambda)}(\gamma'') \end{aligned} \quad (16)$$

where the noncentrality parameter is

$$\lambda = \frac{\tilde{\mathbf{s}}^H \tilde{\mathbf{s}}}{1/2}$$

A comparison of this model to the classical linear model is given in Figure 6. This figure was generated using $\lambda = 20$, $M = 11$ and $N = 32$.

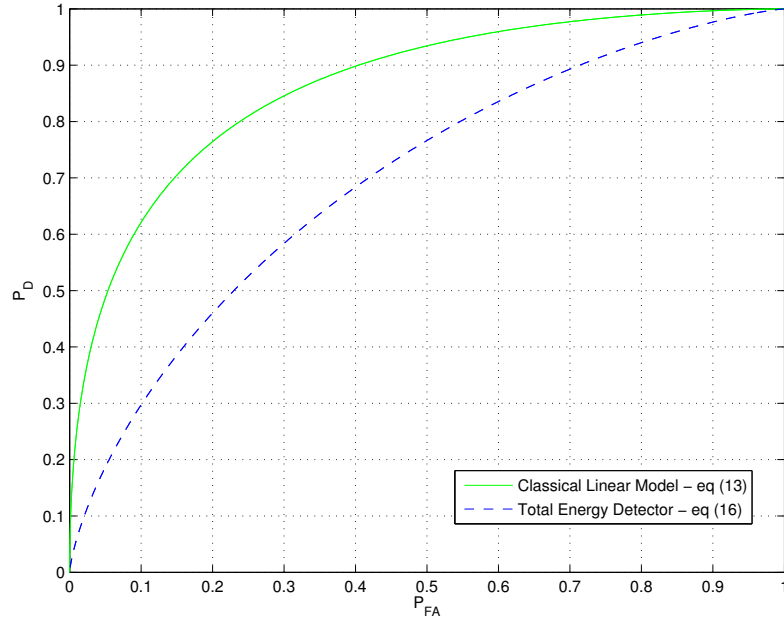


Figure 6. Exact detection performance of the Classical Linear Model compared against the Total Energy Detector.

In situations where the possible target location region is quite large, we have noticed that this detector performs almost as well as the GLRT derived in Section 2.3. This is because the total energy detector, unlike the GLRT, does not depend

on the target location estimate. When the possible target location region is large, using a similar explanation given in [11], it can be shown that the P_{FA} increases with the number of “bins” searched for target location. This translates as a deterioration of the performance of the GLRT. While the total energy detector does not have this degradation, it cannot be used for target localization like the GLRT. Figure 7 shows the comparison of the performance of the GLRT against the total energy detector where we restricted the grid search for the maximum eigenvalue to only a region of $3\text{km} \times 3\text{km}$ square around the actual target location.

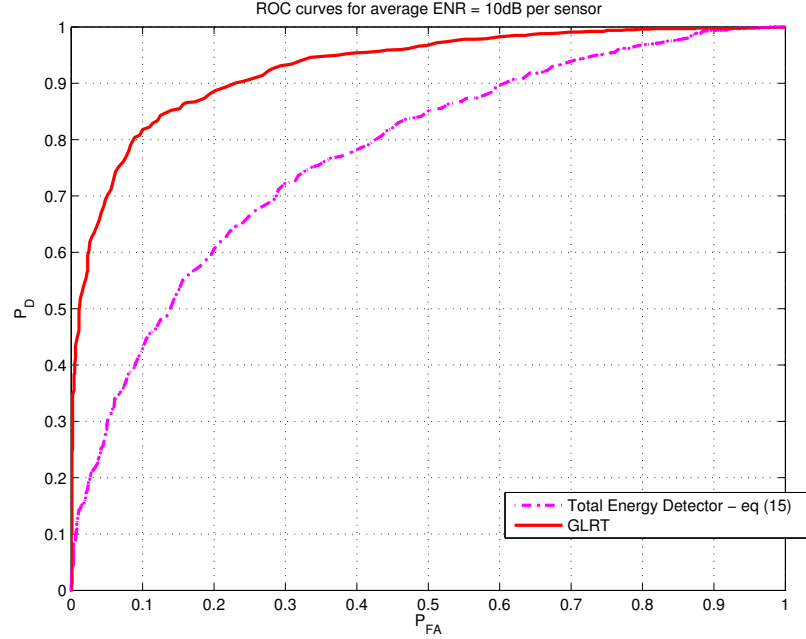


Figure 7. Comparison of GLRT against the Total Energy Detector for observations collected at 11 sensors with unknown time delays and Doppler shifts

2.6 Conclusion

We have modeled the problem of detection of LPI signals and derived the GLRT detector for the model. We have shown that the GLRT uses both the energy and the cross-sensor correlation and thus outperforms the currently used detectors which only use either the energy or the cross correlation. While finding the GLRT, the MLE of the target location and velocity are also simultaneously computed. We also derived simpler detectors by making various assumptions in the main model. We have shown that when a target has to be detected in a large region, the total energy detector is only slightly poorer than the GLRT. In the next paper we will discuss the performance of the MLE of the target location mentioned here. We are also investigating the possibility of extending this technique to the detection of multiple targets.

3 Localization

3.1 Introduction

Passive localization has been used for many years and has always been an important topic of research [17, 18, 19, 20]. Localization can be performed using one or more of the emitter location dependent properties of the signal such as angle of arrival, TDOA, FDOA or the energy of the received signal. Over the years the general approach to localization using the TDOAs, commonly referred to as the TDOA technique, has been to first estimate the difference in the times of arrival of the signal at a particular pair of sensors and then use these TDOAs to estimate the location of the emitter. Knapp and Carter [21, 22] proposed a generalized correlation method for the estimation of the time delay for stationary and relative motion cases. They modeled the signal as a stationary Gaussian random process. Stein [9] on the other hand modeled the signal as deterministic but unknown and derived the MLE for the differential delay and Doppler for a two sensor case. Under similar assumptions for the signal, Yeredor and Angel [23] have derived the CRLB for the TDOAs. After the TDOAs are estimated they are used to estimate the location of the source [24, 25, 26, 27]. Quite often, due to network capacity and computational constraints, not all sensor pair combinations are used. Fowler [28] addressed the problem of optimal selection of a subset of the sensor pairs. Torrieri [17] proposed a linear least squares estimator where the nonlinear relation between the TDOAs and the emitter location is linearized by expanding it in a Taylor series about a reference point and retaining the first two terms. This is an iterative method which requires some kind of a priori information in order to obtain an initial guess. Alternatively, Chan and Ho [18] use an intermediate variable, which is a function of the emitter location, in order to linearize the nonlinear equations. They use a two-step weighted least squares (WLS) algorithm. Additionally, when the signal waveform is known, localization may be performed from the times of arrival (TOAs) instead of the TDOAs [29, 30]. In such TOA based techniques the unknown transmission time occurs as a nuisance parameter which will have to be estimated. Do et. al [31] have shown that the TOA and the TDOA measurements are transformable to each other without a loss of information regarding positioning and thus the position estimations based on them should be theoretically equivalent. The above techniques may be called two-step techniques because the TOAs/TDOAs are first estimated at the local sensors and these TOA/TDOA estimates are used in a second step to compute the location of the emitter.

Weiss and Amar [32, 33, 34, 35] have shown that the two-step approach is sub-optimal and proposed a direct position determination (DPD) approach. Weiss had derived the MLE for the source location for the case of a stationary narrow-band radio frequency transmitter using multiple stationary receivers in [32]. He uses a continuous time model and quickly considers the sampled version without

discussing the effects of sampling on the emitter location estimate. He shows that the MLE of the emitter location is obtained by maximizing a quadratic form of the signal samples whose coefficients are functions of the emitter location. He considers the two cases of signal known and signal unknown but leaves out a more important case - signal known but transmission time unknown which is most likely to occur in real-world situations. He does not discuss the Cramer-Rao lower bound (CRLB) for this problem. There is an inherent ambiguity in the commonly used model and Weiss uses a constraint on the signal samples to resolve the ambiguity. No discussion is provided on the generality of the constraint as to why it is an appropriate constraint, how it resolves the ambiguity and whether it reduces the performance. In [33] Amar and Weiss extend the approach to a multiple emitters case. In [34] they address the problem of localization using only the Doppler frequency shifts and in [35] they consider the case of a single stationary emitter and moving receivers. In all these cases the results are similar, i.e., the MLE for the emitter location is obtained by maximizing a quadratic form of the signal whose coefficients are functions of the emitter location. The derivation of CRLB is attempted in the later papers but is not sufficiently simplified. The effect of sampling the signal is not discussed in any of the papers. Similar constraints are used to resolve the ambiguity in the following papers but no discussion is provided on the effects of the constraint.

In this paper we consider the case of a single stationary emitter and a network of stationary receive sensors. We address many of the short-comings of [32, 33, 34, 35]. We use a continuous time model and provide a straightforward derivation for the MLE of the emitter location for the two cases of signal waveform known with unknown transmission time and signal waveform unknown with unknown transmission time. Our model is valid for either narrow-band or broadband signals, lowpass or highpass signals. We discuss the effect of sampling the signal on the emitter location estimate. Using simulations, we compare the MLE against a conventional TDOA technique. We show that the variance of the MLE is two to three orders of magnitude lower than the conventional TDOA technique.

A more difficult problem is deriving the CRLB. If the signal waveform is assumed unknown along with the time of arrival (TOA) and the attenuation factor, then the commonly used model has an ambiguity. This ambiguity comes to light when deriving the CRLB. Because all the unknowns in the model cannot be uniquely resolved, the Fisher information matrix (FIM) becomes singular. We address this ambiguity in detail and derive the necessary steps to remove it. Then we derive the non-singular FIM. The inverse of the FIM is the CRLB. CRLB gives the theoretical lower bound on the variance of any unbiased estimator.

An important application of the CRLB is in deriving an optimal sensor configuration. The performance of a location estimator depends on the placement of sensors. A particular configuration of the sensors is called optimal if it optimizes a norm of the FIM. A quite common result [36, 37] is to place the sensors around the emitter in an equi-angular configuration. But when the sensors are geographically constrained the problem becomes much more difficult. We introduce this problem

in Section 4.1 and provide optimal sensor configurations for the three and four sensor cases.

In Section 6.2 we provide a detailed description of the problem. In Section 3.3 we give the CRLBs and the MLEs. Here we analyze the case of signal waveform unknown and the special case of signal waveform known, both cases with an unknown transmission time. In Section 6.4 we use Monte Carlo simulations to compare the performance of the MLE against the conventionally used TDOA technique. We show that at higher signal-to-noise ratios (SNRs) the variance of the MLE approaches the CRLB. In section 4.1 we introduce the problem of optimal sensor configurations and give some results. Conclusions are provided in Section 3.5. Most of the mathematics is provided in the appendices. In Appendix B.1 we derive a compact expression for the FIM. Appendix B.2 has the derivation of the MLE. Appendix B.3 presents the properties of a matrix we use in the model. In Appendix B.4 we discuss the transformation of the parameters and the constraints used in order to remove the ambiguity in the model. In Appendix C we derive the unconstrained optimal sensor configuration and the constrained optimal sensor configurations for the three and four sensor cases.

3.2 Methods, Assumptions, and Procedures

Suppose that a stationary emitter is located at an unknown location (x_T, y_T) and a network of M sensors are located at known locations (x_i, y_i) , $i = 0, 1, \dots, M-1$ as shown in Figure 27. For simplicity we are assuming a two dimensional case. Extension to the three dimensional case is straightforward. The sensors are all

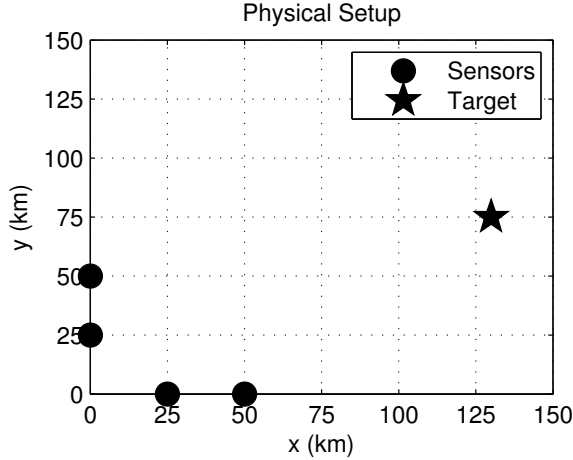


Figure 8. Physical placement of the sensors (for $M=4$) and the Emitter position used for simulation.

synchronized in time and each of the sensors intercepts the signal within the time interval $(0, T)$. The emitter transmits an unknown signal $s(t)$ for an unknown duration $T_s < T$ starting at an unknown time $t_0 < T$. We shall assume that the transmitted signal $s(t)$ is real. It can be narrowband or wideband, lowpass

or bandpass. After interception, the signal received at sensor i in the presence of noise can be written as

$$r_i(t) = A_i s(t - \tau_i) + w_i(t), \quad 0 < t < T, \quad i = 0, 1, \dots, M-1 \quad (17)$$

where $w_i(t)$ is a zero mean wide sense stationary additive white Gaussian random process with spectral density $\frac{N_0}{2}$, A_i 's are the unknown attenuations due to propagation loss, assumed real, and the τ_i 's are the unknown TOAs given by

$$\tau_i = \frac{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}}{c} + t_0, \quad i = 0, 1, \dots, M-1 \quad (18)$$

where c is the propagation speed of the signal. We assume that the noise at a sensor is independent of the noise at any other sensor, i.e., $w_i(t)$ and $w_j(t)$ are independent for $i \neq j$ and that the noise spectral density at all the sensors is equal to $\frac{N_0}{2}$. If the noise does not satisfy these conditions then the problem becomes more complex. For example if the noise spectral density is different at each sensor but known, then the noise term does not factor out as in equation (22) but instead exists in each term. A more difficult problem is when the noise spectral density is different at each sensor and unknown, in which case, the noise spectral densities at each of the sensors need to be estimated as well. To keep the derivations simple we assumed the above conditions for the noise. Notice that here we do not assume as in [9], that $\tau_i \ll T$. Instead we just assume that $\max_{i,j}(\tau_i - \tau_j) < (T - T_s)$. That is, we are only assuming that the observation interval is large enough so that, within the observation interval, the signal reaches both the nearest and the farthest sensors from the emitter. Based on the sensor geometry it is possible to find a sufficient condition on the length of the observation interval. If d_{\max} is the distance between the farthest pair of sensors, then the observation interval must be greater than $\frac{d_{\max}}{c}$.

Sampling the signal in time has a quantization like effect on the estimate of the emitter location. This is because if the signal is sampled, then the TOA estimates are integer multiples of the sampling interval and hence quantized. For example if the signal is sampled at a frequency of F_s samples/sec, then the estimate of τ_i is quantized with a maximum quantization error of $\frac{1}{2F_s}$. This can introduce a maximum quantization error of $\frac{c}{2F_s}$ in the range estimate. Therefore, it is possible that the signal may have to be sampled at a rate much higher than the Nyquist rate in order to achieve a desired precision in the location estimate. We have used a continuous time model to avoid this problem in our analysis and to allow future studies of errors due to time synchronization effects.

Figure 9 shows the signals received at the four sensors shown in Figure 27 from an emitter located at (130, 75)km transmitting a Gaussian chirp. The propagation loss was modeled as a $\frac{1}{R}$ attenuation in the amplitude of the received signal, where R is the range. So, the farthest sensor has the largest TOA and smallest amplitude. We are assuming that the signal lies inside the observation interval $(0, T)$. So, we

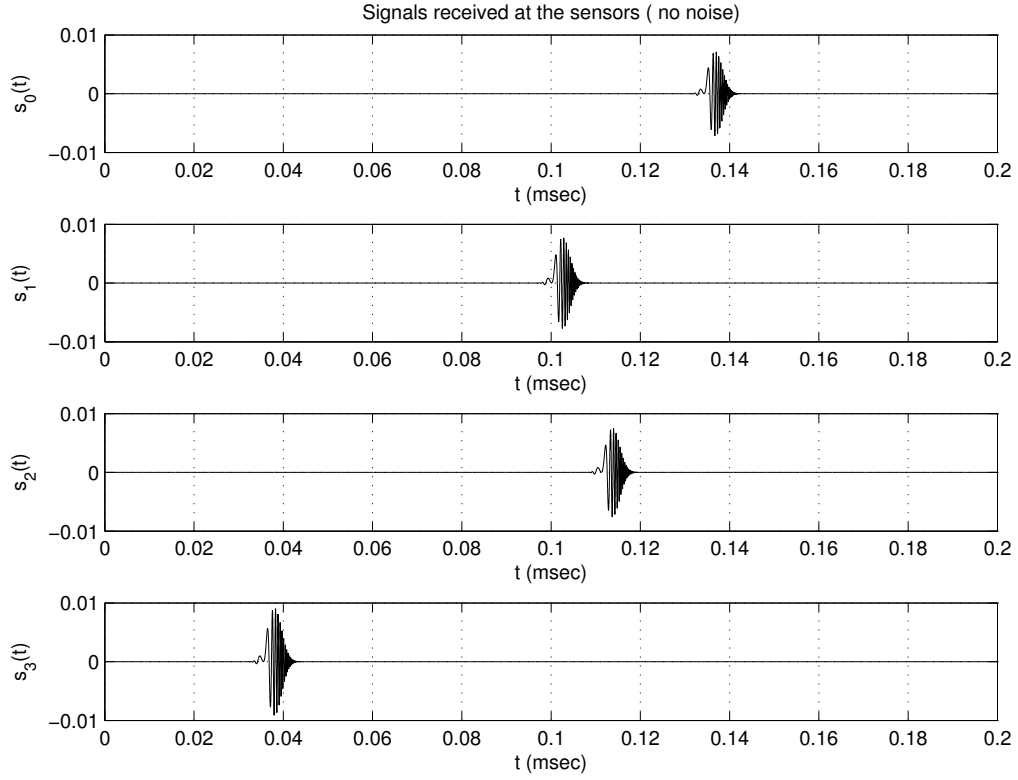


Figure 9. Signals received at the four sensors when a Gaussian chirp is transmitted by the emitter located at (130,75)km.

can assume that the unknown signal is periodic with period T and write it in terms of its Fourier series as

$$s(t) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{\infty} (a_n \cos 2\pi n F_0 t + b_n \sin 2\pi n F_0 t) \quad (19)$$

where $F_0 = \frac{1}{T}$ and the Fourier coefficients are given by

$$a_0 = \frac{\sqrt{2}}{T} \int_0^T s(t) dt, \quad a_n = \frac{2}{T} \int_0^T s(t) \cos 2\pi n F_0 t dt, \quad b_n = \frac{2}{T} \int_0^T s(t) \sin 2\pi n F_0 t dt \quad (20)$$

We are using $\frac{a_0}{\sqrt{2}}$ for the d.c component instead of the standard a_0 because it simplifies certain terms in the derivation of the CRLB. For a band-limited signal only a finite number of the Fourier coefficients are non-zero. If the signal is a lowpass signal, there exists an integer N such that the Fourier coefficients are all zero for $n \geq N$ and if the signal is a bandpass signal, then there exist integers N_1 and N_2 , $N_1 < N_2$, such that the Fourier coefficients are zero for $n < N_1$ and for $n > N_2$. So we can approximate the lowpass signal $s(t)$ as (for a bandpass signal the summation is from N_1 to N_2)

$$s(t) = \frac{a_0}{\sqrt{2}} + \sum_{n=1}^{N-1} (a_n \cos 2\pi n F_0 t + b_n \sin 2\pi n F_0 t)$$

This is an important step as it allows us to model any unknown signal and reduce it to a parameter estimation problem. Now, if we let

$\boldsymbol{\phi} = [a_0 \ a_1 \ \cdots \ a_{N-1} \ b_1 \ b_2 \ \cdots \ b_{N-1}]^T$ be the $2N-1 \times 1$ vector of Fourier coefficients and

$$\mathbf{h}(t) = \left[\frac{1}{\sqrt{2}} \cos 2\pi F_0 t \ \cdots \ \cos 2\pi(N-1)F_0 t \ \sin 2\pi F_0 t \ \cdots \ \sin 2\pi(N-1)F_0 t \right]^T$$

then we have $s(t) = \mathbf{h}^T(t)\boldsymbol{\phi}$. This reduces the uncountable unknown parameter set $\{s(t) : t \in (0, T)\}$ to a finite countable number of unknown parameters $\boldsymbol{\phi}$. Therefore, we can rewrite the model in equation (45) as

$$r_i(t) = A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi} + w_i(t), \quad 0 \leq t \leq T, \quad i = 0, 1, \dots, M-1 \quad (21)$$

Let $\boldsymbol{\tau} = [\tau_0 \ \tau_1 \ \cdots \ \tau_{M-1}]^T$ and $\mathbf{A} = [A_0 \ A_1 \ \cdots \ A_{M-1}]^T$. Let $\boldsymbol{\theta} = [\boldsymbol{\tau}^T \ \mathbf{A}^T \ \boldsymbol{\phi}^T]^T$ be the $(2M+2N-1) \times 1$ vector of unknown parameters. If we let $\boldsymbol{\eta} = [x_T \ y_T \ t_0]^T$ and $\boldsymbol{\alpha} = [\boldsymbol{\eta}^T \ \mathbf{A}^T \ \boldsymbol{\phi}^T]^T$ then, using equation (46), we can write the TOA vector as a function of $\boldsymbol{\eta}$ as $\boldsymbol{\tau} = \mathbf{g}(\boldsymbol{\eta})$. So, the problem can be stated as, given the observations $r_i(t)$, $i = 0, 1, \dots, M-1$ estimate the vector $\boldsymbol{\eta}$. We are only interested in the parameters (x_T, y_T) and the rest of the unknown parameters are nuisance parameters.

3.3 CRLB and MLE of the Emitter Location

Here we will derive the CRLB and the MLE of the emitter location for the two cases of signal waveform unknown with unknown transmission time and signal waveform known with unknown transmission time.

3.3.1 Signal unknown with unknown transmission time

For the continuous time model in equation (45) the log-likelihood function [38] for sensor i is given by

$$l = -\frac{1}{N_0} \int_0^T (r_i(t) - A_i s(t - \tau_i))^2 dt \quad (22)$$

Since the noise at different sensors is independent, and using equation (21) we can write the joint log-likelihood function as

$$l(\boldsymbol{\theta}) = -\frac{1}{N_0} \int_0^T \sum_{i=0}^{M-1} (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi})^2 dt \quad (23)$$

The $(2M+2N-1 \times 2M+2N-1)$ FIM [39] for this model is given by (see Appendix B.1.1)

$$\mathcal{I}_{\boldsymbol{\theta}} = \frac{(T/2)}{(N_0/2)} \begin{bmatrix} (2\pi F_0)^2 \boldsymbol{\phi}^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi} (\text{diag}(\mathbf{A}))^2 & (2\pi F_0)(\boldsymbol{\phi}^T \mathbf{L} \boldsymbol{\phi})(\text{diag}(\mathbf{A})) & (2\pi F_0)(\mathbf{A} \odot \mathbf{A}) \boldsymbol{\phi}^T \mathbf{L} \\ (2\pi F_0)(\boldsymbol{\phi}^T \mathbf{L}^T \boldsymbol{\phi})(\text{diag}(\mathbf{A})) & (\boldsymbol{\phi}^T \boldsymbol{\phi}) \mathbf{I}_M & \mathbf{A} \boldsymbol{\phi}^T \\ (2\pi F_0) \mathbf{L}^T \boldsymbol{\phi} (\mathbf{A} \odot \mathbf{A})^T & \boldsymbol{\phi} \mathbf{A}^T & (\mathbf{A}^T \mathbf{A}) \mathbf{I}_M \end{bmatrix} \quad (24)$$

where \odot represents the element by element product (Hadamard product), $\text{diag}(\mathbf{A})$ is an $M \times M$ diagonal matrix with i th diagonal element as A_i , \mathbf{I}_M is the $M \times M$ identity matrix and the $(2N - 1) \times (2N - 1)$ matrix \mathbf{L} is given by

$$\mathbf{L} = \begin{bmatrix} & \mathbf{0}_{(N,N)} & \begin{bmatrix} \mathbf{0}_{(1,N-1)} \\ \text{diag}(1, 2, \dots, N-1) \end{bmatrix} \\ - \begin{bmatrix} \mathbf{0}_{(N-1,1)} & \text{diag}(1, 2, \dots, N-1) \end{bmatrix} & & \mathbf{0}_{(N-1,N-1)} \end{bmatrix}$$

This FIM must be inverted in order to find the CRLB for the unknown parameter vector $\boldsymbol{\theta}$. By the form of the matrix in equation (24) it is easily shown (see Appendix B.4) that *the matrix is singular with rank equal to two less than full rank*. Weiss [32] uses an ad hoc method to overcome this. We, however use the exact transformation of the parameters [40] that is required to eliminate the singularity of the information matrix. The singularity arises because there is an ambiguity in the model. It is not possible to uniquely determine all the unknown parameters in the model in equation (45). This is because of the relationship between the TOA, attenuation factor and the signal waveform. Suppose that in equation (45), \bar{A}_i and $\bar{s}(t)$ are the true values of the gain and the signal waveform that generate $r_i(t)$. Then the pair of values $(\bar{A}_i/\gamma, \gamma\bar{s}(t))$ for any non-zero constant γ also generate the same $r_i(t)$. So, from the observation $r_i(t)$ it is impossible to determine the true values of A_i and $s(t)$. A similar relationship exists between the TOA and the signal waveform. Suppose that in equation (45), $\bar{\tau}_i$ and $\bar{s}(t)$ are the true values of the TOA and the signal waveform that generate $r_i(t)$. Then the pair of values $((\bar{\tau}_i - \gamma), \bar{s}(t - \gamma))$ for any constant γ also generate the same $r_i(t)$. This is more clearly demonstrated in Figure 10. Notice that given the received signal it is not possible to determine whether the source signal is $s_1(t)$ with TOA τ_1 or $s_2(t)$ with TOA τ_2 . This causes the information matrix to be at least rank two

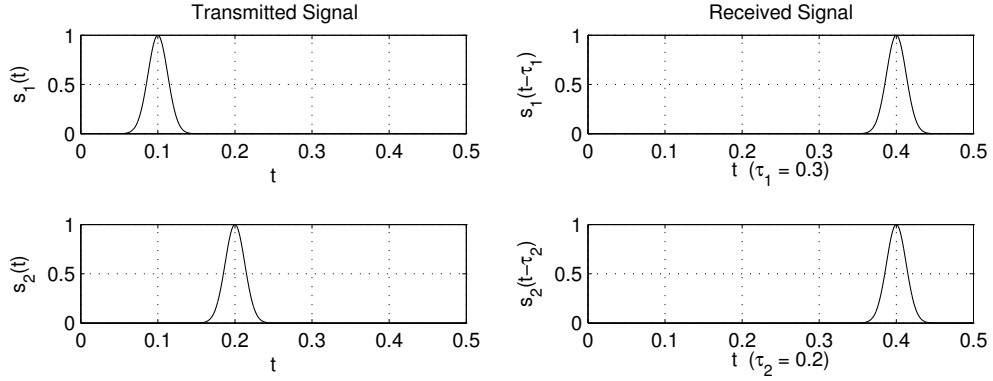


Figure 10. Ambiguity when signal and TOA are both unknown.

deficient, as shown in Appendix B.4. The over parameterization can be resolved by applying an appropriate transformation that satisfies certain constraints [40].

As shown in Appendix B.4, the appropriate transformation for this model is

$$\begin{aligned}\boldsymbol{\tau}' &= [(\tau_1 - \tau_0) \quad (\tau_2 - \tau_0) \quad \cdots \quad (\tau_{M-1} - \tau_0)]^T \\ \mathbf{A}' &= (1/A_0)[A_1 \quad \cdots \quad A_{M-1}]^T \\ \boldsymbol{\phi}' &= A_0 \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}(\mathbf{h}(-\tau_0))\boldsymbol{\phi}\end{aligned}\quad (25)$$

and can be shown to be the least restrictive constraint for identifiability. Here we are using the $(*)'$ notation to represent the new parameters resulting from the transformation. Notice that the transformed parameter vectors $\boldsymbol{\tau}'$ and \mathbf{A}' are TDOA and relative gain factor with respect to sensor 0 and are each reduced by one parameter from $\boldsymbol{\tau}$ and \mathbf{A} respectively while the $\boldsymbol{\phi}'$ is simply the Fourier coefficients of $r_0(t)$. Using these transformed parameters the model in equation (21) can be rewritten as

$$\begin{aligned}r_0(t) &= \mathbf{h}^T(t)\boldsymbol{\phi}' + w_0(t) & 0 \leq t \leq T \\ r_i(t) &= A'_i \mathbf{h}^T(t - \tau'_i)\boldsymbol{\phi}' + w_i(t), & 0 \leq t \leq T, \quad i = 1, 2, \dots, M-1\end{aligned}\quad (26)$$

where $A'_i = \frac{A_i}{A_0}$ and $\tau'_i = \tau_i - \tau_0$. So, the effect of the transformation is that the signal at sensor 0 is made the reference signal and the signals at all the other sensors are modeled relative to this reference signal. Although (26) seems intuitively obvious, by arriving at it from (21) using the transformation in (25), we have mathematically verified that (26) is indeed the correct model to use for the problem of localization under the unknown signal case. Weiss [32, 33, 34, 35] uses (26) directly without this rigorous argument. This is a subtle but important result which is overlooked by Weiss. Now, let $\boldsymbol{\theta}' = [\boldsymbol{\tau}'^T \quad \mathbf{A}'^T \quad \boldsymbol{\phi}'^T]^T$ be the $(2M + 2N - 3) \times 1$ vector of the unknown transformed parameters, $\boldsymbol{\eta}' = [x_T, y_T]^T$ and $\boldsymbol{\alpha}' = [\boldsymbol{\eta}'^T \quad \mathbf{A}'^T \quad \boldsymbol{\phi}'^T]^T$. Notice that the TDOA vector $\boldsymbol{\tau}'$ is a function of only (x_T, y_T) , i.e. $\boldsymbol{\tau}' = \mathbf{g}'(\boldsymbol{\eta}')$. The unknown transmission time t_0 does not appear and thus is not a nuisance parameter. The problem is now, given the observations $r_i(t)$, $i = 0, 1, \dots, M-1$ estimate the vector $\boldsymbol{\eta}'$. The log-likelihood function for this model with the transformed parameters is given by

$$l(\boldsymbol{\theta}') = -\frac{1}{N_0} \int_0^T (r_0(t) - \mathbf{h}^T(t)\boldsymbol{\phi}')^2 dt - \frac{1}{N_0} \int_0^T \sum_{i=1}^{M-1} (r_i(t) - A'_i \mathbf{h}^T(t - \tau'_i)\boldsymbol{\phi}')^2 dt \quad (27)$$

As shown in the Appendix B.1.1, the FIM for this transformed parameter vector is

$$\mathcal{I}_{\boldsymbol{\theta}'}^{-1} = \mathbf{H} \mathcal{I}_{\boldsymbol{\theta}}^{\dagger} \mathbf{H}^T \quad (28)$$

where \mathbf{H} is given in (B.29), or equivalently,

$$\mathcal{I}_{\boldsymbol{\theta}'} = \frac{T/2}{(N_0/2)} \begin{bmatrix} (2\pi F_0)^2 \boldsymbol{\phi}'^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi}' (\text{diag}(\mathbf{A}'))^2 & (2\pi F_0)(\boldsymbol{\phi}'^T \mathbf{L} \boldsymbol{\phi}') (\text{diag}(\mathbf{A}')) & (2\pi F_0)(\mathbf{A}' \odot \mathbf{A}') \boldsymbol{\phi}'^T \mathbf{L} \\ (2\pi F_0)(\boldsymbol{\phi}'^T \mathbf{L}^T \boldsymbol{\phi}') (\text{diag}(\mathbf{A}')) & (\boldsymbol{\phi}'^T \boldsymbol{\phi}') \mathbf{I}_{(M-1)} & \mathbf{A}' \boldsymbol{\phi}'^T \\ (2\pi F_0) \mathbf{L}^T \boldsymbol{\phi}' (\mathbf{A}' \odot \mathbf{A}')^T & \boldsymbol{\phi}' \mathbf{A}'^T & (\mathbf{1} + \mathbf{A}'^T \mathbf{A}') \mathbf{I}_{(2N-1)} \end{bmatrix}. \quad (29)$$

The FIM for the corresponding vector $\boldsymbol{\alpha}'$ is given by [39]

$$\mathcal{I}_{\boldsymbol{\alpha}'} = \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T} \right)^T \mathcal{I}_{\boldsymbol{\theta}'} \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T} \right) = \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T} \right)^T (\mathbf{H} \mathcal{I}_{\boldsymbol{\theta}}^\dagger \mathbf{H}^T)^{-1} \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T} \right) \quad (30)$$

where the Jacobian $\left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T} \right)$ is given in equation (B.18). In Appendix B.2.1 we show that, the MLE for the emitter location is obtained by maximizing over (x_T, y_T) , the maximum eigenvalue of the $M \times M$ cross-correlation matrix $\mathbf{B}' = \mathbf{Y}' \mathbf{Y}'^T = \sum_{i=0}^{M-1} \mathbf{y}'_i \mathbf{y}'_i{}^T$ where $\mathbf{Y}' = [\mathbf{y}'_0 \ \mathbf{y}'_1 \ \cdots \ \mathbf{y}'_{M-1}]$ with

$$\mathbf{y}'_0 = \int_0^T r_0(t) \mathbf{h}(t) dt \quad \text{and} \quad \mathbf{y}'_i = \int_0^T r_i(t) \mathbf{h}(t - \tau'_i) dt, i = 1, 2, \dots, M-1. \quad (31)$$

That is,

$$\hat{\boldsymbol{\eta}}' = \arg \max_{\boldsymbol{\eta}'} \lambda_{\max}(\mathbf{B}') \quad (32)$$

or equivalently,

$$(\hat{x}_T, \hat{y}_T) = \arg \max_{(x_T, y_T)} \lambda_{\max}(\mathbf{B}') \quad (33)$$

where λ_{\max} represents the maximum eigenvalue. The matrix \mathbf{B}' is real symmetric and positive definite and so the maximum eigenvalue is real and positive.

3.3.2 Signal known with unknown transmission time

Quite often in practical situations it is possible that the signal waveform is known but the exact transmission time t_0 is unknown. In this case the number of unknowns is reduced to $2M$. Let $\boldsymbol{\zeta} = [\boldsymbol{\tau}^T \ \mathbf{A}^T]^T$ be the $2M \times 1$ unknown parameter vector. Similar to (23), the log-likelihood function is given by

$$l(\boldsymbol{\zeta}) = -\frac{1}{N_0} \int_0^T \sum_{i=0}^{M-1} (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi})^2 dt \quad (34)$$

where $\boldsymbol{\phi}$ is known. The FIM for this model is given by (see Appendix B.1.2)

$$\mathcal{I}_{\boldsymbol{\zeta}} = \frac{(T/2)}{(N_0/2)} \begin{bmatrix} (2\pi F_0)^2 \boldsymbol{\phi}^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi} (\text{diag}(\mathbf{A}))^2 & (2\pi F_0) (\boldsymbol{\phi}^T \mathbf{L} \boldsymbol{\phi}) (\text{diag}(\mathbf{A})) \\ (2\pi F_0) (\boldsymbol{\phi}^T \mathbf{L}^T \boldsymbol{\phi}) (\text{diag}(\mathbf{A})) & (\boldsymbol{\phi}^T \boldsymbol{\phi}) \mathbf{I}_M \end{bmatrix} \quad (35)$$

This matrix is not singular because, for this case, the unknown parameters in the model can be uniquely determined. Therefore, there is no need to transform the parameters as in the case of the unknown signal. Although, the unknown transmission time t_0 is still retained here as the nuisance parameter. For this model it is shown in Appendix B.2.2 that the MLE for emitter location and the unknown transmission time is given by

$$\hat{\boldsymbol{\eta}} = \arg \max_{\boldsymbol{\eta}} \boldsymbol{\phi}^T \mathbf{B} \boldsymbol{\phi} \quad (36)$$

where $\mathbf{B} = \mathbf{Y}\mathbf{Y}^T$ and $\mathbf{Y} = [\mathbf{y}_0 \ \mathbf{y}_1 \ \cdots \ \mathbf{y}_{M-1}]$ with $\mathbf{y}_i = \int_0^T r_i(t) \mathbf{h}(t - \tau_i) dt$, $i = 0, 1, \dots, M-1$. \mathbf{B} is a function of (x_T, y_T, t_0) . Using the fact that $\mathbf{h}^T(t - \tau_i) \boldsymbol{\phi} = s(t - \tau_i)$ and $\boldsymbol{\eta} = [x_T \ y_T \ t_0]^T$ we can rewrite equation (36) as

$$(\hat{x}_T, \hat{y}_T, \hat{t}_0) = \arg \max_{(x_T, y_T, t_0)} \sum_{i=0}^{M-1} \left(\int_0^T r_i(t) s(t - \tau_i) dt \right)^2 \quad (37)$$

Equation (37) is simply the correlation values between the known signal waveform and the observed signal at each of the sensors, summed over all the sensors. The emitter location that yields the values of the TOAs that maximize the expression in equation (37) is the MLE of the emitter location.

3.4 Results and Discussion

In order to evaluate the performance of the MLE we have run some simulations and compared the performance against the CRLB and against a typically used TDOA approach. The TDOA approach was implemented as a two-step algorithm where, in the first step, the TDOA $\Delta \hat{\tau}_i$ estimates were obtained by cross-correlating the signal at each sensor with the signal at sensor 0. Then a $1\text{km} \times 1\text{km}$ region around the true emitter location was split into 100×100 grid points and for each emitter location on the grid point the TDOAs were computed using the formula

$$\Delta \tau_i = \frac{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}}{c} - \frac{\sqrt{(x_T - x_0)^2 + (y_T - y_0)^2}}{c}, \quad i = 1, 2, \dots, M-1$$

Next the least squared error (LSE) between the estimated TDOAs and the computed TDOAs was calculated as

$$\text{LSE} = \sum_{i=1}^{M-1} (\Delta \hat{\tau}_i - \Delta \tau_i)^2$$

This LSE is a function of the emitter location (x_T, y_T) . The emitter location that minimized the LSE is the estimate of the emitter location. Since, we know that the LSE is a 2-dimensional parabolic function of the emitter location, we improved the accuracy by fitting a parabola through the 10000 points at which the LSE was computed. Then by using the analytical formula for the minimum location of a 2-dimensional parabola, we computed the minimum. To measure the levels of the zero mean additive white Gaussian noise, we used a metric called the average SNR (ASNR). The ASNR is the ratio of the average signal power to the noise power at each sensor averaged over all the sensors, i.e., if $\mathcal{P}_{si} = |A_i|^2 \frac{1}{T} \int_0^T |s(t)|^2 dt$ is the average power of the signal at the i th sensor and $\frac{N_0}{2}$ is the noise spectral density at the i th sensor, then the SNR averaged over M sensors is given by $10 \log \left(\frac{1}{M} \sum_{i=0}^{M-1} \frac{\mathcal{P}_{si}}{(N_0/2)(F_s/2)} \right)$ dB. Figure 11 (a) shows a realization of the LSE function. For the simulation we have used 4 sensors placed at the

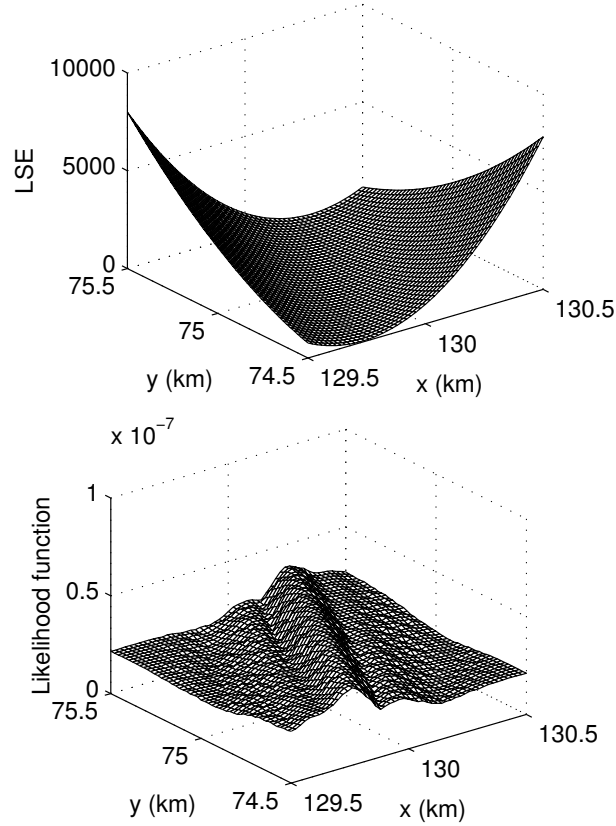


Figure 11. (a) A realization of the LSE function at an ASNR = -10 dB. (b) A realization of the Likelihood function at an ASNR = -10 dB.

coordinates shown in Figure 27 and the emitter was placed at the coordinates (130, 75) km, also as shown in Figure 27. The integrations in (20) and (31) are approximated using summations with $\delta t = 0.33$ ns which means the sampling frequency is $F_s = 300$ MHz. The reason for choosing such high sampling frequency is that at this frequency, the position quantization error due to sampling is in the order of $(c/F_s) = 10^{-3}$ km. A Gaussian chirp defined by

$$s(t) = \exp\left(-\frac{1}{2}\sigma_F^2\left(t - \frac{T_s}{2}\right)^2\right) \sin(2\pi mt^2)$$

was used as the unknown transmitted signal waveform. Figure 12 shows the transmitted signal waveform. Notice that the signal is assumed to be approximately zero for $t < 0$ and for $t > T_s$. We set $T_s = 5\mu\text{s}$ and $\sigma_F = 200,000\pi$. The observation interval at each of the sensors was taken to be $T = 0.2$ ms. The unknown transmission time of the signal was set to $t_0 = 0.07$ ms. With this configuration the maximum TDOA is 0.0988 ms. The frequency spectrum of the Gaussian window is given by $|S(F)| = (\sqrt{2\pi}/\sigma_F) \exp(-2\pi^2 F^2/\sigma_F^2)$ [39], and the bandwidth of the chirp is $BW = 1.5$ MHz. The rate of change of frequency for the linear chirp was

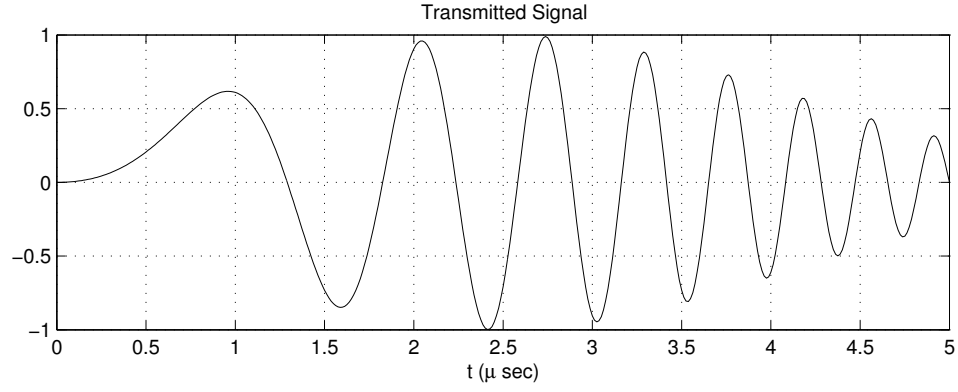


Figure 12. Transmitted signal waveform.

chosen to be $m = \frac{BW}{T_s} = 3 \times 10^8$. A plot of the Fourier coefficients of the signal is shown in Figure 13. A value of $N = 2 \times BW \times T = 600$ was used to have a total of $2N - 1 = 1199$ unknown Fourier coefficients. Notice that the Fourier coefficients are almost zero for $n > 600$. We set the ASNR at -20 dB and ran a total of 300

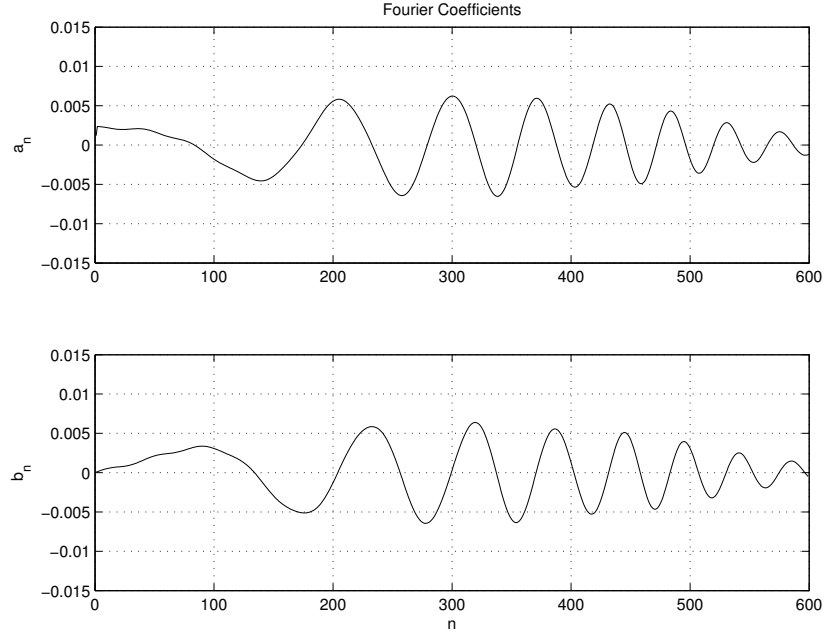


Figure 13. Fourier coefficients plot.

Monte Carlo simulations to generate the scatter plot and the corresponding 95% error ellipse which are shown in Figure 14. This is also called the 95% confidence ellipse. That is, if this estimator is used a large number of times for localization, then around 95% of those times the true location of the emitter will lie within this ellipse. To compute the MLE the maximization was performed using a grid search.

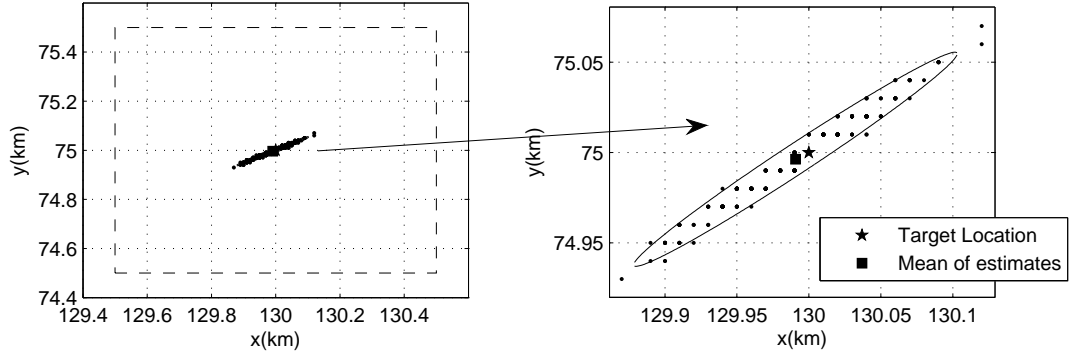


Figure 14. Scatter plot and the corresponding 95% error ellipse of the MLE for an ASNR = -20 dB.

We used a grid of size $1 \text{ km} \times 1 \text{ km}$ with the grid points 0.01 km apart to have a total of $100 \times 100 = 10000$ points. The grid is shown with a dotted line in the figure. At -20 dB , the variances of the MLEs of (x_T, y_T) were $(0.0021, 0.0006) \text{ km}^2$ and the respective CRLBs were $(0.0012, 0.0003) \text{ km}^2$. The figure does not show 300 points because some points lie on top of the others due to position quantization induced by the finite number of grid points in the grid search for maximization. Due to the complex nature of the likelihood function (see Figure 11 (b)), it is not possible to use any curve fitting techniques to reduce this quantization effect as in the case of the conventional TDOA approach.

Figure 15 shows the comparison of the variances of the MLE and the typical TDOA approach against the CRLB for different SNR values. Notice that for the ASNR values below -30 dB , the variance of the MLE remains flat. This is because of the restriction imposed by the finite grid size. As the ASNR increases above -30 dB , the variance of the MLE reduces rapidly to approach the CRLB at around -10 dB . Due to the nature of the conventional TDOA approach it breaks down for the ASNR values below -17 dB . So this figure has the variances of the TDOA approach only for the average SNR values above -17 dB . On the other hand, for this particular setup, the results for the MLE are reliable for the ASNR values as low as -30 dB . It is quite obvious from this figure that performance of the MLE is very much better than a typical TDOA approach. In this case, the MLE performs as good as a typical TDOA approach for an ASNR value of about 10 dB less than that for the TDOA approach. Also notice that at around -10 dB , the variance of the MLE is almost two orders of magnitude less than that of the TDOA approach. We have noticed that for certain sensor-emitter configurations, particularly when the emitter was very close to the sensors, the variance of the MLE is up to three orders of magnitude less than that of the TDOA approach. Therefore, under LPI scenarios where a conventional TDOA technique cannot be reliably used, the MLE can be used.

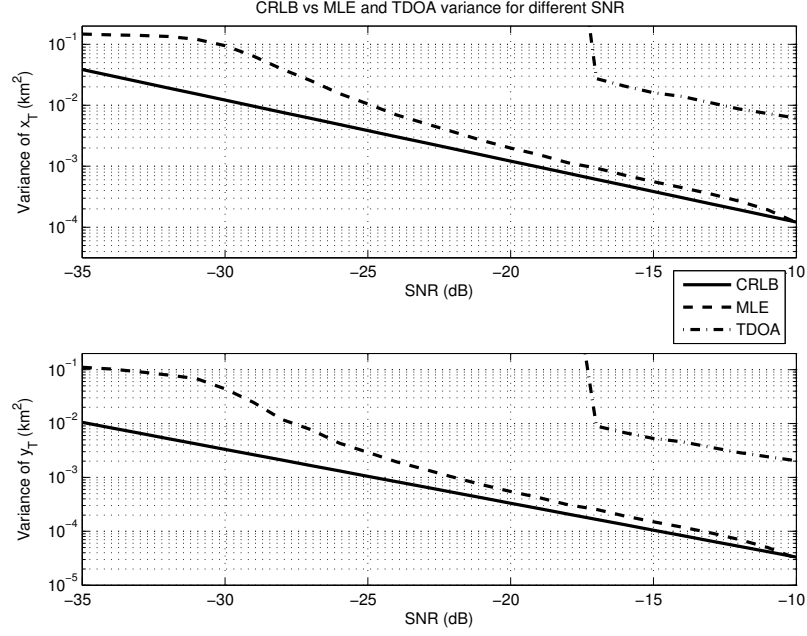


Figure 15. Comparison of variances for the emitter location estimate using MLE and a typical TDOA approach against the CRLB for different average SNR values.

3.5 Conclusions

We have derived a direct positioning estimator for an emitter location. This is the maximum likelihood estimator. We have shown that for an unknown signal case, the model that is conventionally used has an inherent ambiguity and so all the unknown parameters cannot be uniquely determined. We derived an appropriate transformation of the parameters and re-parameterized the model to remove the ambiguity. We have shown that for the special case of a known signal with unknown transmission time, there is no ambiguity in the model. We derived the MLE and the FIM for the model. The performance of the MLE was compared against a typical two-step TDOA based localizer and against the CRLB. The performance of the MLE is significantly better than a typical two-step TDOA based localizer.

4 Optimal Sensor Configuration

4.1 Introduction

Apart from helping evaluate the performance of the MLE, the CRLB has other important applications. One such application is to determine the optimal sensor configurations. From equations (30) and (B.18) we see that the FIM of the emitter location vector $[x_T \ y_T]^T$ depends on the Jacobian $\left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T}\right)$. Now, this Jacobian is dependent on the sensor emitter geometry. Figure 16 shows different geometries and the corresponding CRLBs for an ASNR of -20 dB. It can be noticed that for configurations that surround the emitter, the CRLB is smaller. The problem of

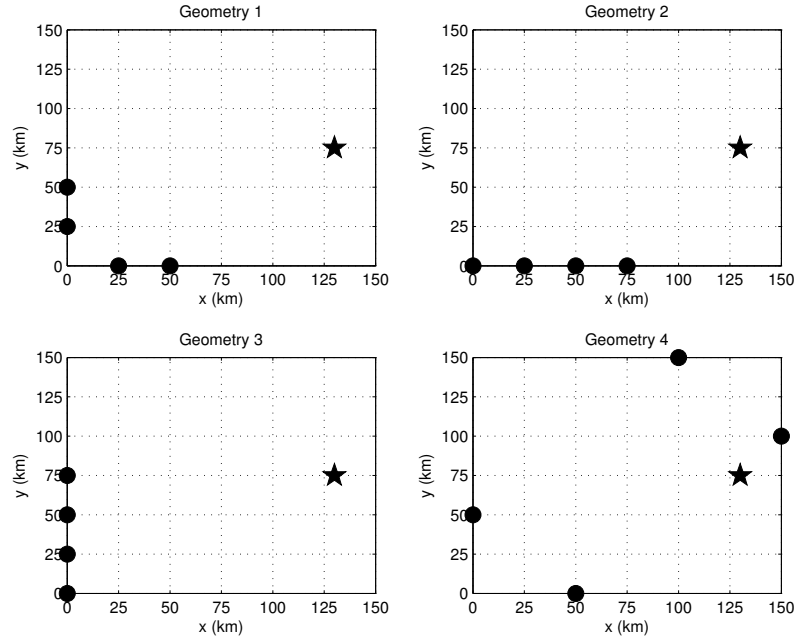


Figure 16. Different Sensor Configurations.

optimizing the sensor configuration falls under a broader class of problems called “Optimal Design of Experiments”. Generally, a norm of the FIM is chosen as the optimization criterion and depending on the norm, the corresponding configura-

Table 2. CRLB for different geometries

Geometry	CRLB(x)	CRLB(y)
1	0.0012	0.0006
2	0.0024	0.0021
3	0.0015	0.0001
4	0.0656×10^{-4}	0.2852×10^{-4}

tion is classified as A-optimal, where the trace of the inverse of the information matrix is minimized, D-optimal, where the determinant of the information matrix is maximized, T-optimal, where the trace of the information matrix is maximized etc. Lee [36] has derived the A-optimal sensor configurations for the conventional TDOA approach. He has shown that the optimal configuration is to place the sensors with an equi-angular spacing around the emitter. Similar results were derived by Bishop et al [37] for the D-optimal configuration. We can notice consistent results in the example in Figure 16. When the sensors are around the emitter as in Case 4, the CRLB is greatly reduced. Placing the sensors around the emitter may not however be practically feasible. A practical problem of interest is to find an optimal sensor configuration where the sensors are restricted to a sector. The A-optimal configuration for this problem is addressed by Lee [36].

4.2 Methods, Assumptions, and Procedures

Now, from (C.1), we have the FIM for the emitter location as

$$\mathcal{I}_{\eta'} = \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T (2\pi F_0)^2 \boldsymbol{\phi}'^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi}' (\text{diag}(\mathbf{A}'))^2 \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)$$

If we assume that the signal level is the same at all sensors, ($A_i = A$) for all i , we have the above FIM as

$$\mathcal{I}_{\eta'} = (2\pi F_0)^2 \boldsymbol{\phi}'^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi}' A^2 \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)$$

The geometry dependent term in the above equation is $\left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)$. Therefore, an emitter-sensor configuration that maximizes the determinant of $\left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)$ is called the D-optimal configuration.

We define the unconstrained problem as one where the sensors are allowed to be anywhere in angle (equivalent to being on a circle) and, in the constrained configuration problem, the sensors can only lie on an arc of the circle.

4.3 Results and Discussion

We have derived the D-optimal configuration for three and four sensor cases (see Appendix C). That is, we derived an emitter-sensor configuration that maximizes the determinant of the FIM of the emitter location. In Appendix C.1 we show that for the *unconstrained problem*, a D-optimal configuration is to place the sensors at equi-angular distance around the emitter. This appears rather contradictory since it is required to know the location of the emitter in order to optimally configure the sensors for localization. In fact, this result has applications in mobile emitter localization where in the recursive tracking of the emitter using a sequence of TDOA measurements is performed [41]. This result is consistent with

the results of Lee and Bishop. Our proof in Appendix C.1 is much simpler than previously published results. But this configuration is not the unique D-optimal configuration. As the number of sensors increases, there are many more configurations which are D-optimal but are difficult to find. For example, in the case of six sensors, placing the sensors at an angular distance of $\frac{\pi}{3}$ radians from each other on the circle around the emitter, as shown in Figure 17 (a), is one D-optimal configuration. Another D-optimal configuration is to place two sensors at each of the vertices of an equilateral triangle that is inscribed in the circle around the emitter, as shown in Figure 17 (b). Here the sensors are supposed to lie on top of each other but for visual clarity, they are slightly displaced.

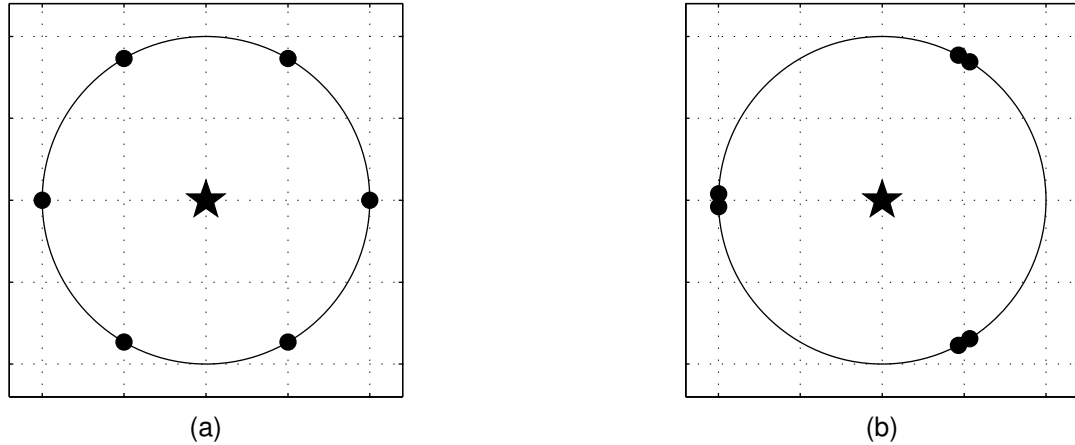


Figure 17. Two D-optimal configurations for the case of $M=6$ sensors.

The constrained problem is much more difficult to solve. The setup of a constrained problem is as shown in Figure 18. The sensors can lie only on the arc with half angle θ which is marked in bold in the figure. Using the transformation

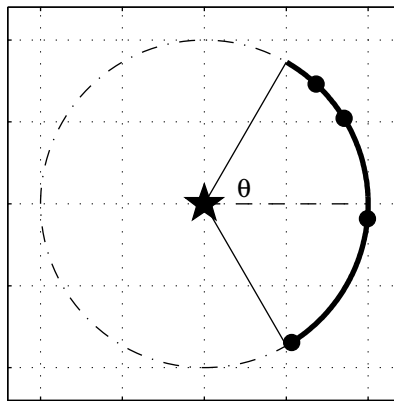


Figure 18. Constrained sensor configuration setup. The sensors can lie only on the arc marked in bold.

of variables given in Appendix C we have the determinant as

$$\left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) = \det(\mathbf{J}\mathbf{J}^T)$$

where

$$\mathbf{J} = \begin{bmatrix} \cos \psi_0 & \cos \psi_1 & \cdots & \cos \psi_{M-1} \\ \sin \psi_0 & \sin \psi_1 & \cdots & \sin \psi_{M-1} \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

where ψ_i is the angle of sensor i from the positive x-axis measured at the emitter and $-\theta \leq \psi_i \leq \theta$ for all $i = 0, \dots, M-1$. Notice that when $M = 3$, we have

$$\mathbf{J} = \begin{bmatrix} \cos \psi_0 & \cos \psi_1 & \cos \psi_2 \\ \sin \psi_0 & \sin \psi_1 & \sin \psi_2 \\ 1 & 1 & 1 \end{bmatrix}$$

and $\det(\mathbf{J}\mathbf{J}^T)$ is proportional to the square of the area of the triangle with vertices at the sensor locations. So, in geometrical terms the problem can be viewed as finding the triangle with the maximum area that has vertices on the constraining arc. In order to increase the area, the base and the height of the triangle must be increased. So, intuitively it can be seen that the optimal configuration for the three sensor case is as shown in Figure 19 (A rigorous proof is provided in Appendix C.2).

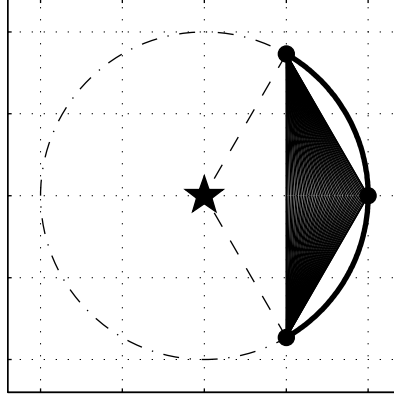


Figure 19. The D-optimal configuration for three sensors. The triangle with the maximum area is also shown here.

This geometrical interpretation can be extended to any number of sensors using the Cauchy-Binet formula. For an arbitrary M , we have

$$\det(\mathbf{J}\mathbf{J}^T) = \sum_{S \in \binom{[M]}{3}} \det(\mathbf{J}_S \mathbf{J}_S^T)$$

where $[M]$ is the set $\{0, 1, \dots, M-1\}$, and $\binom{[M]}{3}$ is the set of 3-combinations of $[M]$ (i.e., subsets of size 3). \mathbf{J}_S is the 3×3 matrix whose columns are the columns

of \mathbf{J} at indices from S . Now, each $\det(\mathbf{J}_S \mathbf{J}_S^T)$ is the square of the area of the triangle formed by the sensors with indices in S . So, maximizing the determinant is equivalent to maximizing the sum of the squares of the areas of the triangles formed by all combinations of sensors taken three at a time. There is a total of $\binom{M}{3}$ such combinations. We hypothesize that in general a D-optimal configuration for any number of sensors will consist of distributing the sensors in appropriate proportions at the center and at the two end points of the arc. There are three D-optimal configurations for the four sensor case, which are given in Figure 20 (proof in Appendix C.2). These derivations are very specific to the particular number of sensors and are hard to extend to higher number of sensors. In a future paper we will discuss this constrained optimal configuration in detail.

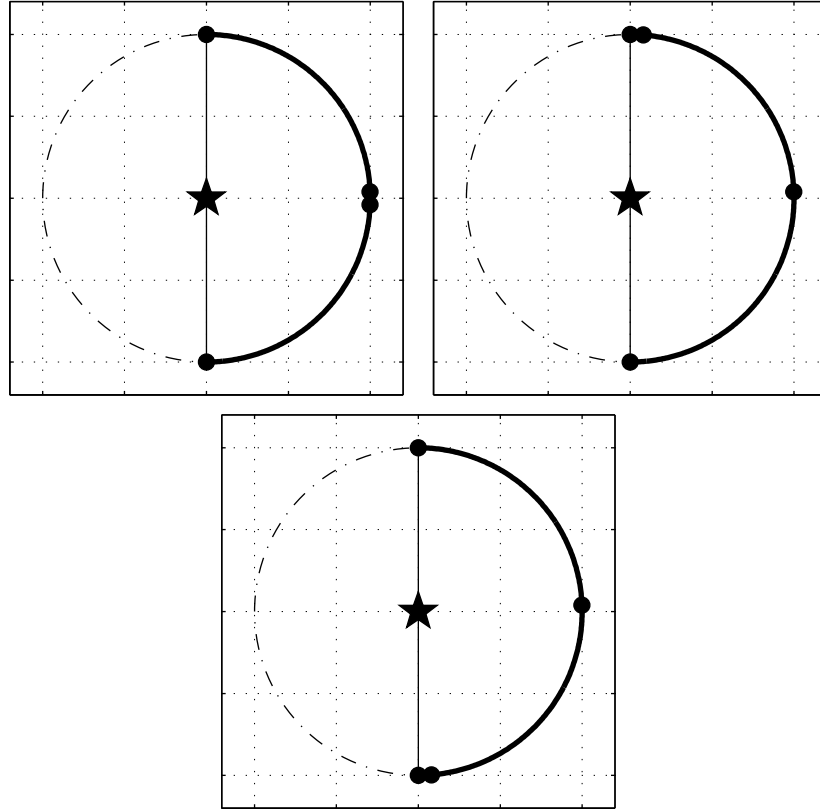


Figure 20. The three D-optimal configurations for four sensor case.

4.4 Conclusions

We have introduced the problem of optimizing the sensor-emitter geometry. We derived the D-optimal configuration for the unconstrained geometry problem. For the constrained geometry problem, we have derived the D-optimal configurations for the three and four sensor cases.

5 Knowledge Aided Design

5.1 Introduction

In knowledge aided design, the knowledge of the terrain is used for localization. Presence of objects in a terrain induces some azimuth modulation due to obstruction, reflection, diffraction etc of the signal. Knowledge of the location of these obstacles and the type of modulation they induce, can aid in the localization of the target of interest. This is particularly useful in localization in the urban environments where there is buildings, trees etc which act as the obstacles. In this section we will derive the Fisher information matrix and show how the azimuth modulation induced by an obstacle can increase the information.

5.2 Methods, Assumptions, and Procedures

Suppose that we have M sensors located at (x_i, y_i) $i = 1, \dots, M$. Let an emitter located at an unknown location (x_T, y_T) transmit a known signal waveform $s(t)$. Assuming that the sensors are in a direct line of sight from the emitter and that there is no multipath, the signal received at the sensors is modeled as

$$r_i(t) = A_i s(t - \tau_i) + w_i(t) \quad i = 0, \dots, M - 1 \quad (38)$$

where A_i s are the unknown attenuation factors and τ_i s are the unknown time-delays. $w_i(t)$ is white Gaussian noise. We want to determine the amount of information in the azimuth modulation.

5.3 Results and Discussion

The FIM for the model in (38) is given by (see Appendix D)

$$\begin{aligned} \mathcal{I}(x_T, y_T) = & \frac{\mathcal{E}}{(N_0/2)} \frac{\bar{F}^2}{c^2} \sum_{i=1}^M A_i^2(x_T, y_T) \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} + \\ & \frac{\mathcal{E}}{(N_0/2)} \sum_{i=1}^M \begin{bmatrix} \left(\frac{\partial A_i}{\partial x_T} \right)^2 & \left(\frac{\partial A_i}{\partial x_T} \right) \left(\frac{\partial A_i}{\partial y_T} \right) \\ \left(\frac{\partial A_i}{\partial y_T} \right) \left(\frac{\partial A_i}{\partial x_T} \right) & \left(\frac{\partial A_i}{\partial y_T} \right)^2 \end{bmatrix} \end{aligned} \quad (39)$$

Here the A_i s and τ_i s are known functions of the target location. For example, assuming there are no obstacles, we have

$$A_i(x_T, y_T) = \frac{G_i}{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}} = \frac{G_i}{R_i}$$

where G_i is a constant and

$$\tau_i(x_T, y_T) = (1/c) \sqrt{(x_T - x_i)^2 + (y_T - y_i)^2} = (1/c) R_i$$

where c is the propagation speed of the signal. So, the FIM for this problem is given by (see Appendix D.1)

$$\begin{aligned}\mathcal{I}(x_T, y_T) &= \frac{\mathcal{E}}{(N_0/2)} \sum_{i=1}^M \left\{ \frac{\bar{F}^2}{c^2} \cdot \frac{G_i^2}{R_i^2} \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} + \right. \\ &\quad \left. \frac{G_i^2}{R_i^4} \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} \right\} \\ &= \mathcal{I}_\tau + \mathcal{I}_A\end{aligned}\quad (40)$$

where the first term is the contribution from the time-delays and the second term is the contribution from azimuth modulation. Note that the information in the azimuth modulation is greater than the information in the time-delays if

$$\frac{\bar{F}^2}{c^2} \cdot \frac{G_i^2}{R_i^2} < \frac{G_i^2}{R_i^4} \quad \Rightarrow \quad R_i < \sqrt{\frac{c^2}{\bar{F}^2}}$$

That is, for short ranges as in urban environments, or for signals with narrow band-width, the information in the azimuth modulation becomes comparable to the information in the time-delays. A possible range and bandwidth is given in the following example.

5.3.1 Example 1 - No Obstacles

We have 4 sensors and an emitter located at (945, 810)m as shown in Figure 21. The transmitted signal is a Gaussian pulse given by

$$s(t) = \exp \left(-\frac{1}{2} \sigma_F^2 \left(t - \frac{T_s}{2} \right)^2 \right)$$

The length of the signal in time is $T_s = 50\mu s$ and its root-mean-square (RMS) bandwidth is $\sqrt{\bar{F}^2} = 111.07$ kHz. This is also shown in Figure 21. So we have the energy of the signal $\mathcal{E} = 11.284 \times 10^{-6}$. The noise spectral density $N_0 = 2.2568 \times 10^{-6}$ so that the energy to noise ratio is $\frac{\mathcal{E}}{(N_0/2)} = 10$. As an example assume, the gain at the sensors is taken as $G = 60$ dB. For this setup we have the FIM as

$$\mathcal{I} = \mathcal{I}_\tau + \mathcal{I}_A = \begin{bmatrix} 23.7315 & 17.8134 \\ 17.8134 & 20.1021 \end{bmatrix}$$

where

$$\mathcal{I}_\tau = \begin{bmatrix} 2.9743 & 2.1807 \\ 2.1807 & 2.3558 \end{bmatrix} \quad \text{and} \quad \mathcal{I}_A = \begin{bmatrix} 20.7572 & 15.6327 \\ 15.6327 & 17.7463 \end{bmatrix} \quad (41)$$

Now, with the same setup but with a signal of RMS bandwidth $\sqrt{\bar{F}^2} = 5.55$ MHz the FIMs were

$$\mathcal{I} = \begin{bmatrix} 7456.6 & 5467.3 \\ 5467.3 & 5907.3 \end{bmatrix}$$

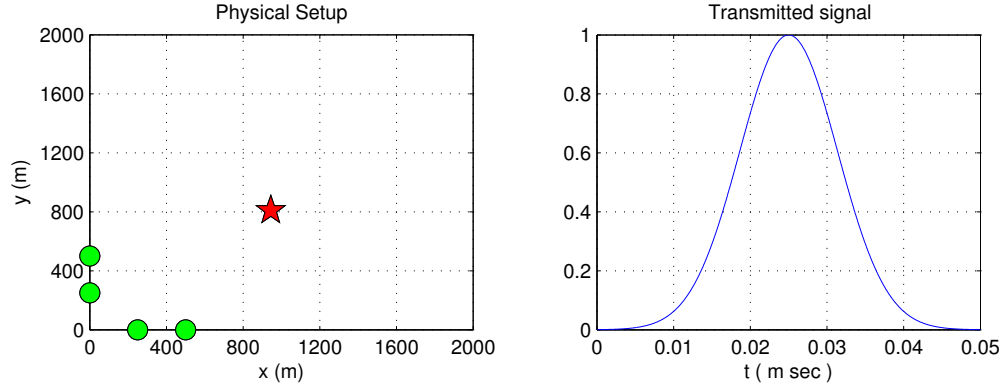


Figure 21. Physical setup and Transmitted signal

where

$$\mathcal{I}_\tau = \begin{bmatrix} 7435.9 & 5451.7 \\ 5451.7 & 5889.5 \end{bmatrix} \quad \text{and} \quad \mathcal{I}_A = \begin{bmatrix} 20.7572 & 15.6327 \\ 15.6327 & 17.7463 \end{bmatrix} \quad (42)$$

As the bandwidth of the signal increased, the information in the time-delays increased but the information in the azimuth modulation remained the same. \square

Therefore, for signals with RMS bandwidth in the order of 100 kHz and ranges in the order of a couple of kilometers, the information in the azimuth modulation is comparable to the information in the time-delays. For larger distance or for signals with larger bandwidths, the information in the time-delays far exceeds the information in the azimuth modulation and so even if there is any increase in azimuth modulation due to presence of buildings it may not significantly increase the total information.

5.3.2 With an Obstacle

Next, suppose that there is an obstacle. Some of the receivers may not have a direct line of sight to the emitter. To analyze one aspect of the obstacles, viz. blocking of the signal, let us assume that they do not reflect the signal. So, there is no multipath and so the signal received at the sensors is still given by (38). But now the A_i s are affected by the location of the obstacles. For simplicity assume that there is one obstacle at (x_B, y_B) . This obstacle induces some kind of azimuth modulation, say $f_B(x_T, y_T)$ so that we have

$$A_i(x_T, y_T) = \left(\frac{G_i}{R_i} \right) f_B(x_T, y_T)$$

Now, the information matrix is given by (see Appendix)

$$\mathcal{I}(x_T, y_T) = \frac{\varepsilon}{(N_0/2)} \sum_{i=1}^M \left\{ \left(\frac{\bar{F}^2}{c^2} \cdot \frac{G_i^2}{R_i^2} + \frac{G_i^2}{R_i^4} \right) f_B^2(x_T, y_T) \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} + \left(\frac{G_i^2}{R_i^4} \right) f_B^2(x_T, y_T) \mathbf{J}_i \right\} \quad (43)$$

where the matrix \mathbf{J}_i is given in (D.3). Compare this to (40) and notice that the information matrix is now scaled by a factor of $f_B^2(x_T, y_T)$. This factor is usually less than one since the obstacles do not induce any gain. Also, notice that there is an additional term $\left(\frac{G_i^2}{R_i^4} \right) f_B^2(x_T, y_T) \mathbf{J}_i$. This term is almost zero everywhere except around the lines joining the obstacle to each of the sensors. This can be seen in Figure 22. This is a plot of the azimuth modulation at the sensor at (250, 0)m (magenta color) as a function of the emitter location. Notice that as the emitter moves away from the line joining the sensor and the obstacle, the attenuation decreases. The derivatives of the azimuth modulation function around these lines are significant and so the contribution of the \mathbf{J}_i matrix to the total information matrix increases.

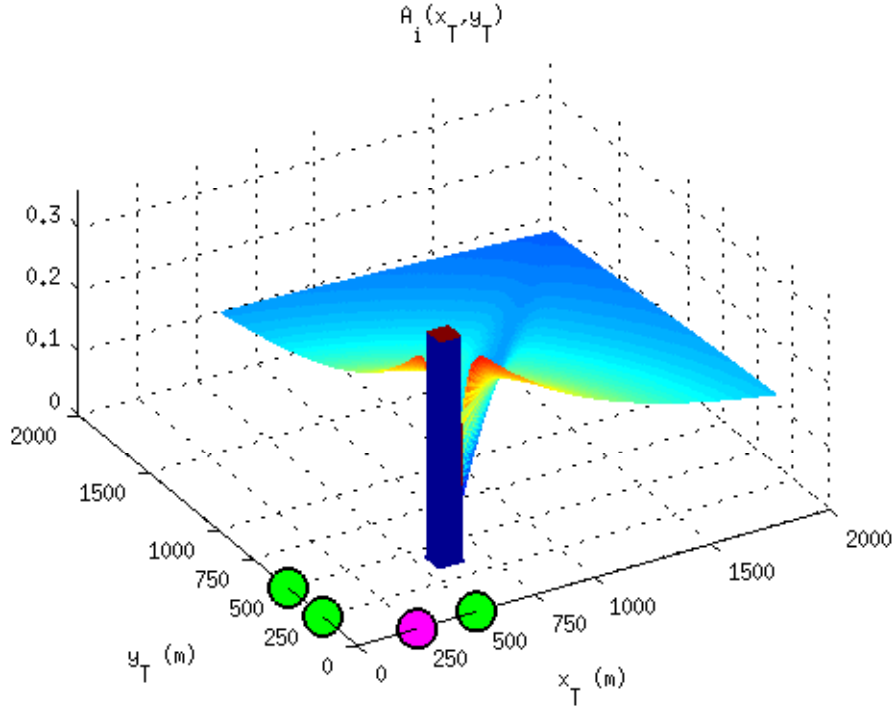


Figure 22. The azimuth modulation function, $A_i(x_T, y_T)$ from (44), at the sensor at (250, 0)m when there is an obstacle at (600, 400)m.

5.3.3 Example 2 - A Building as an Obstacle

Now suppose that there is a building at (600, 400)m as shown in Figure 23. The building is going to block the signal if it is exactly between the emitter and

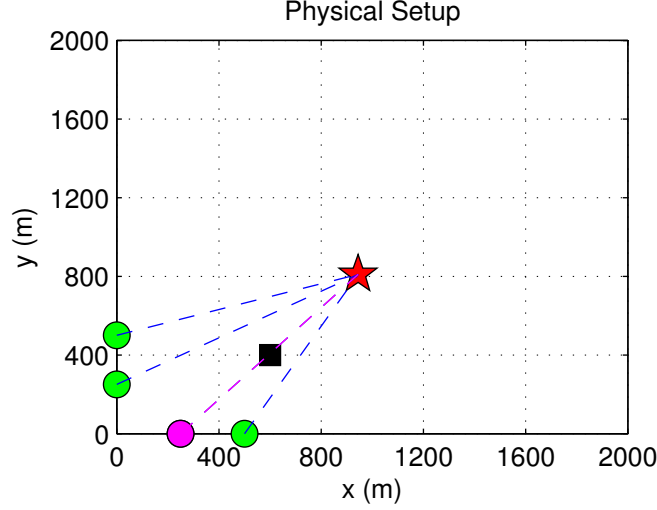


Figure 23. Setup with obstacle very close to the line of sight from the emitter to the sensor at (250, 0)m

the sensor. We model the blocking of the signal by the building as a Gaussian pulse. i.e the signal at a sensor is completely blocked when the building is exactly between the sensor and the emitter and the attenuation loss reduces as a Gaussian pulse on either side perpendicular to the direct line of sight. The attenuation also reduces as the distance of the emitter from the building increases. So we have,

$$A_i(x_T, y_T) = \left(\frac{G}{R_i} \right) \left[1 - \exp \left(-\frac{1}{2} \mathbf{z}^T (\mathbf{Q}^T \mathbf{C}^{-1} \mathbf{Q}) \mathbf{z} \right) \right] \quad (44)$$

where $\mathbf{z} = [(x_T - x_B) \ (y_T - y_B)]^T$. The matrix \mathbf{C} and the rotation matrix \mathbf{Q} are given by

$$\mathbf{C} = \begin{bmatrix} \sigma_W^2 & 0 \\ 0 & \sigma_D^2 \end{bmatrix} \quad \mathbf{Q} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}$$

where σ_W^2 and σ_D^2 determine the width of the building and how the attenuation changes with distance from the building. The angle θ is the angle of the line joining the building and the sensor i from the positive x-axis which is given by $\theta = \tan^{-1} \left(\frac{y_i - y_B}{x_i - x_B} \right)$. The rotation matrix is required to orient the major axis of the Gaussian bell along the line joining the building and the sensor. Figure 22 shows this azimuth modulation at the sensor at (250, 0)m as a function of the target position when a building is at (600, 400)m.

Using the same setup as in Example 1 but now with the building placed at (600, 400)m as in Figure 23 there was a significant increase in the FIM. The new

FIM was (compare this to (41))

$$\mathcal{I} = \begin{bmatrix} 299.4020 & -352.6595 \\ -352.6595 & 416.8441 \end{bmatrix}$$

where

$$\mathcal{I}_\tau = \begin{bmatrix} 0.3661 & 0.2684 \\ 0.2684 & 0.2900 \end{bmatrix} \quad \text{and} \quad \mathcal{I}_A = \begin{bmatrix} 299.0359 & -352.9279 \\ -352.9279 & 416.5541 \end{bmatrix}$$

Figure 24 shows the localization ellipses for the two information matrices. The green represents the localization ellipse when the building is not present and the red represents the localization ellipse when the building is present. Notice that the minor axis is considerably reduced while the major axis is elongated. This is because when the building blocks the signal to a sensor, we know that the emitter is on the line joining the building and the sensor. Here the gain G was reduced to 5000 to make the ellipses fit in the figure.

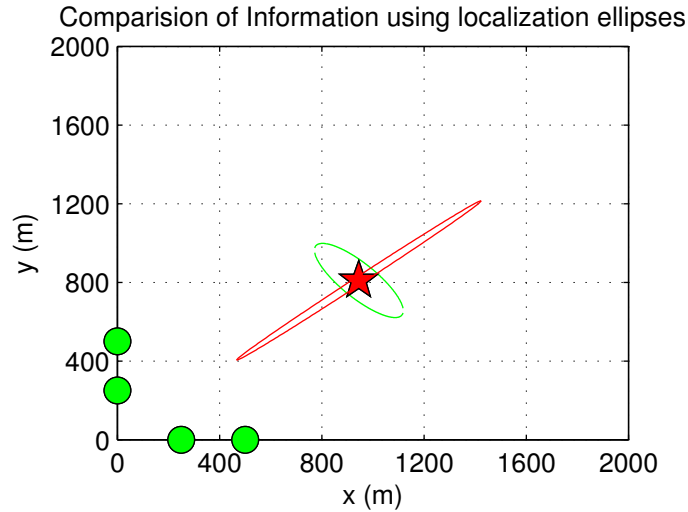


Figure 24. Localization ellipses for the information matrices. Green - no building. Red - With building.

Now, when the building is placed at the point (420, 400)m which is away from the lines joining the emitter to the sensors as shown in Figure 25, the information matrix was not affected and remained the same as in Example 1.

Another interesting fact is that, for the azimuth modulation function we have assumed in (44), the derivative at any point exactly on the line joining the sensor and building is less than at any point on either side of the line. So, when we place the building exactly on that line at (600, 407.9137)m, the information matrix reduced to

$$\mathcal{I} = \begin{bmatrix} 7.2817 & 9.0970 \\ 9.0970 & 12.4407 \end{bmatrix}$$

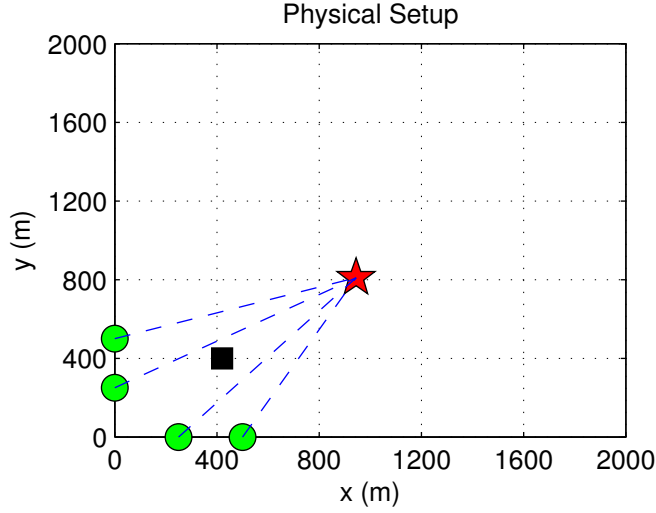


Figure 25. Setup with the obstacle away from the lines of sight from the emitter to the sensors.

Next, with the same setup but with a signal of RMS bandwidth $\sqrt{F^2} = 5.55$ MHz the FIMs were (compare this to (42))

$$\mathcal{I} = \begin{bmatrix} 1214.4 & 318.2 \\ 318.2 & 1141.5 \end{bmatrix}$$

where

$$\mathcal{I}_\tau = \begin{bmatrix} 915.3456 & 671.0949 \\ 671.0949 & 724.9918 \end{bmatrix} \quad \text{and} \quad \mathcal{I}_A = \begin{bmatrix} 299.0359 & -352.9279 \\ -352.9279 & 416.5541 \end{bmatrix}$$

Figure 26 shows the corresponding localization ellipses. So, the total information is reduced. Even though the building increased the information in the azimuth modulation it reduced the information in the time-delays resulting in a net decrease in the total information. \square

5.4 Conclusions

Under certain specific circumstances having the building can increase the Fisher information matrix if the building is close to the lines of sight from the emitter to the sensors. This is because, when the signal at any one sensor is considerably attenuated then we know that the emitter is on the line joining the building and the sensor. This narrows down the possible emitter position from a plane to a single line. Then the signals received at the other sensors are useful in determining the precise location of the emitter on that line. When the building is not close to the direct path between the sensor and the emitter then there is no increase in information due to the presence of the building.

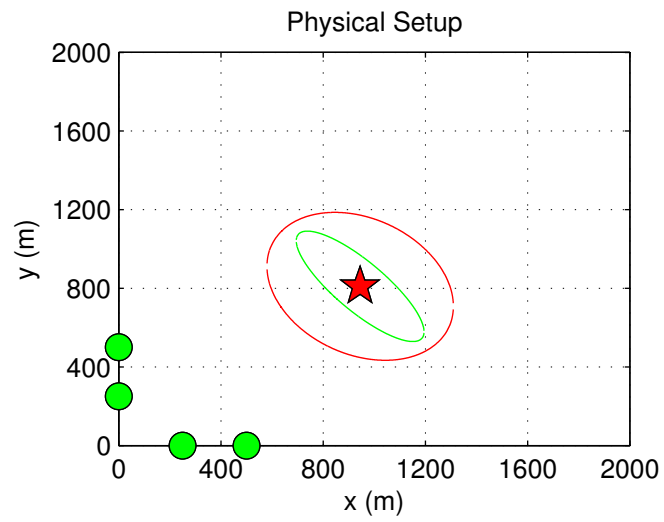


Figure 26. Localization ellipses for the information matrices. Green - no building. Red - With building.

6 Improved TDOA Position Fixing

6.1 Introduction

Localization is the process of extracting the location information of a signal transmitter by intercepting the transmitted signal. The location information can be obtained from one or more of the received signal properties such as signal strength, angle of arrival, TDOA etc. In this paper we focus on the TDOA based localization, also called multilateration. A conventional multilateration has two steps. In the first step the TDOAs between at least four pairs of sensors are estimated. In the second step, these TDOAs are used along with the known sensor locations for position fixing. Abundant literature exists focusing on one or both of these steps. Much of the original work, particularly in sonar, modeled the transmitted signal as a stochastic process with the power spectrum known, in active localization or unknown, in passive localization [25, 42, 24, 21]. The noise at each sensor is almost always assumed to have a known power spectrum and uncorrelated from sensor to sensor. For this model and assumptions, Hahn and Tretter [42] proposed a delay vector estimator using correlators. They showed that when there are $M(> 2)$ sensors, even though we are only interested in the $M - 1$ TDOAs with respect to a reference sensor, it is beneficial to compute the $M(M - 1)/2$ TDOAs for all possible sensor pairs combinations using cross-correlation and then use the Gauss-Markov estimate of the desired $M - 1$ TDOAs with respect to the reference sensor. They also showed that such TDOA estimates attain the CRLB. Knapp and Carter [21] on the other hand proposed a generalized correlation method for the estimation of the TDOAs where in, they pre-filter the received signals in order to maximize the SNR before correlating. Fowler and Hu have shown that the above model and assumptions are reasonable for sonar but not for radar [28]. They have showed that, for passive localization using radars, the signal must be modeled as deterministic and unknown. For this signal model assumption Stein [9] has derived the MLE for the TDOA between two sensors. Stein's result, similar to Hahn and Tretter's [42], is that the MLE for the TDOA between two sensors is the differential delay that maximizes the cross-correlation of the signals received at the two sensors.

Irrespective of the signal model assumptions and how the TDOAs are estimated, each TDOA between a pair of sensors "fixes" the position of the emitter on a hyperboloid with foci at the locations of the two sensors. In 2-dimensions, a set of three TDOAs uniquely define the location of the emitter as the single intersection point of their respective hyperbolas. Therefore, position fixing using the TDOAs is a non-linear problem. Torrieri [17] proposed a linear least squares estimator where this nonlinear relation between the TDOAs and the emitter location is linearized by expanding it using Taylor series about a reference point and retaining the first two terms. Ho and Chan [43, 18] derived a closed form two-step WLS estimator that is asymptotically efficient. They first introduce an intermediate variable that

is a function of the emitter location to linearize the nonlinear equations and find the least squares (LS) estimator. Next, they use the relation between the emitter location and the intermediate variable to solve a second WLS to arrive at the final solution. Recently, Matthew and Silverman [20] have derived a simple closed-form least square estimator that performs quite well.

Alternately, when the signal waveform is known, as in active localization for example, the received signals are correlated with the known signal waveform to estimate the TOAs [39]. Here, it is common to work with the TOAs instead of the TDOAs [29, 30, 44]. The TOA at a sensor “fixes” the emitter position on a sphere (circle in two dimensions) centered at the sensor location. Here, the intersection point of the spheres is taken as the estimate of the emitter location. In the TOA based estimation the unknown transmission time often occurs as a nuisance parameter that needs to be estimated as well. Although, Do et. al [31] have shown that the TOA and TDOA measurements are transformable to each other without loss of information regarding positioning and thus the position estimators based on them should be theoretically equivalent.

These position fixing techniques are good, even efficient, in using the information in the TDOAs. But there is more information in the received signals than just the TDOAs. For example, a measure of the accuracy of each of the TDOAs can also be obtained. This measure can be used to weight the TDOAs so that the more accurate TDOA estimates have a larger contribution in the position fixing step. Such WLS type techniques have also been proposed in the literature. But often the weighting matrix is the inverse of the covariance matrix of the TDOAs which is conveniently assumed to be known [27, 19]. In some cases where the SNR is known, it is used for the weighting.

Recently, Yeredor and Angel [23] have proposed a method where the FIM is used as the weighting matrix. In order to compute the FIM, the emitter location must be known. So they use an iterative technique where they perform estimation over consecutive observation intervals and use the emitter location estimated with the previous observation interval for computing the FIM to be used in the current observation interval. Also, the CRLB (the inverse the FIM) is “on the average” the asymptotic variance of the MLE. So, it is not a very accurate measure of the quality of the TDOA estimate from that particular observation. In this paper we propose a novel approach to estimating the covariance matrix simultaneously with the TDOAs from a given observation interval and then use it for the position fixing step to improve the localization performance.

6.2 Methods, Assumptions, and Procedures

Our focus is on radar and so we use the signal modeling assumptions of Stein [9]. Suppose that a stationary emitter is located at an unknown location (x_T, y_T) and a network of M sensors are located at known locations (x_i, y_i) , $i = 0, 1, \dots, M - 1$ as shown in Figure 27. For simplicity we are assuming a two dimensional case. Extension to the three dimensional case is straightforward. The

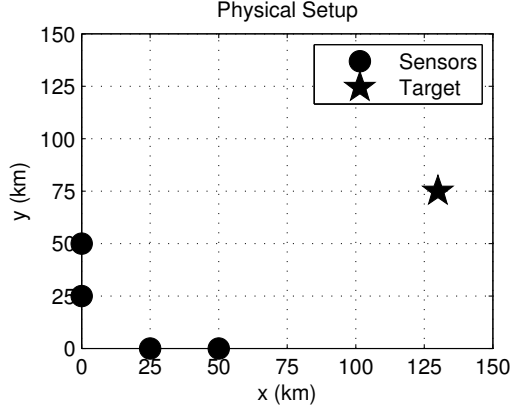


Figure 27. Physical placement of the sensors (for $M=4$) and the Emitter position used for simulation.

sensors are all synchronized in time and each of the sensors intercepts the signal within the time interval $(0, T)$. The emitter transmits an unknown signal $s(t)$ for an unknown duration $T_s < T$ starting at an unknown time $t_0 < T$. We shall assume that the transmitted signal $s(t)$ is real. It can be narrowband or wideband, lowpass or bandpass. After interception, the signal received at sensor i in the presence of noise can be written as

$$r_i(t) = A_i s(t - \tau_i) + w_i(t), \quad 0 < t < T, \quad i = 0, \dots, M-1 \quad (45)$$

where $w_i(t)$ is a zero mean wide sense stationary additive white Gaussian random process, A_i 's are the unknown attenuations due to propagation loss, assumed real, and the τ_i 's are the unknown TOAs given by

$$\tau_i = \frac{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}}{c} + t_0, \quad i = 0, \dots, M-1 \quad (46)$$

where c is the propagation speed of the signal. We assume that the noise at a sensor is independent of the noise at any other sensor, i.e., $w_i(t)$ and $w_j(t)$ are independent for $i \neq j$.

6.3 Variance of the TDOAs

Let the vector of the TOAs be

$$\boldsymbol{\tau} = [\tau_0 \quad \dots \quad \tau_{M-1}]^T \quad (47)$$

and the vector of TDOAs with respect to sensor 0 be

$$\boldsymbol{\tau}' = [\tau'_1 \quad \dots \quad \tau'_{M-1}]^T \quad (48)$$

where $\tau'_i = \tau_i - \tau_0$, $i = 1, \dots, M - 1$ and $(*)^T$ represents the transpose of the matrix $(*)$. Stein has derived the MLE for the TDOA between a pair of sensors. The MLE for τ'_i is given by

$$\hat{\tau}'_i = \arg \max_{\tau \in (-T, T)} \int_0^T r_i(t) r_0(\tau + t) dt. \quad (49)$$

In order to estimate the TDOAs at all sensors with respect to the reference sensor a common practice is to use Stein's MLE to estimate each of the TDOAs individually. That is,

$$\hat{\boldsymbol{\tau}}' = [\hat{\tau}'_1 \quad \dots \quad \hat{\tau}'_{M-1}]. \quad (50)$$

Note that this $(M - 1) \times 1$ TDOA vector estimate in (50) is not the MLE of the TDOA vector in (48). In fact, when $M > 2$, Hahn and Tretter [42] showed that computing the $M(M - 1)/2$ TDOAs for all possible sensor pairs using cross-correlation and then using the Gauss-Markov estimate of the desired $M - 1 \times 1$ vector of TDOAs with respect to the reference sensor is better than (50) and attains the CRLB. When M is large computing all the $M(M - 1)/2$ TDOAs may be impractical. We will continue to use (50) as the estimator for (48) as our focus is on improving the localization accuracy by weighting the TDOAs. Instead, we will further use (49) to also estimate the variances of each of the $\hat{\tau}'_i$ s.

It is a well known fact that the argument that maximizes the likelihood function is the MLE while the asymptotic variance is equal to the negative of the expected value of the curvature of the likelihood function at the peak. *Therefore, we will use the curvature of the likelihood function as the estimate of the inverse of the variance of the corresponding TDOA. This indicates the quality of that particular estimate and not the "on the average" type quality measure as with the CRLB.* The likelihood function in (49) is the correlation function itself given by

$$R(\tau) = \int_0^T r_i(t) r_0(\tau + t) dt. \quad (51)$$

Figure 28 shows segments of two correlation functions around their peak, for different SNRs. Here the SNR at the reference sensor is set at 0 dB. The SNRs at two other sensors that have the same TDOA are set at 0 dB and -10 dB. The signals received at the two sensors are correlated with the signal at the reference sensor. The peak value is subtracted from the corresponding correlation function to bring them to the same level for visual comparison. We define the SNR at each sensor as the average power of the received signal to the average noise power at the sensor, i.e, if at the i th sensor, $\mathcal{P}_{si} = |A_i|^2 \frac{1}{T} \int_0^T |s(t)|^2 dt$ is the average power of the signal and $\frac{N_0}{2}$ is the noise spectral density, then the SNR is given by $10 \log \left(\frac{\mathcal{P}_{si}}{(N_0/2) \times BW} \right)$ dB where BW is the bandwidth of the receiver. Notice that the signal with the higher SNR has a larger curvature indicating a more accurate TDOA and hence should be weighted more heavily. Now, the actual curvature

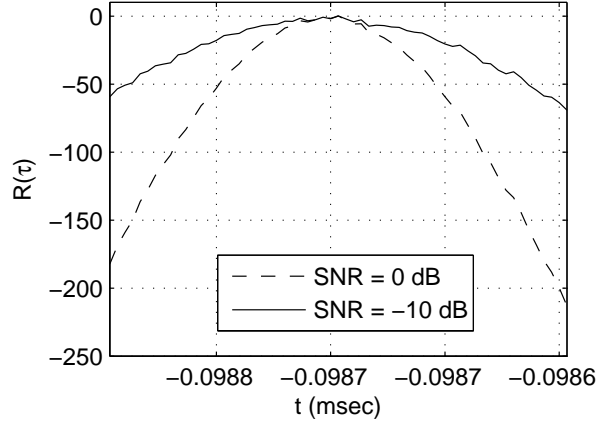


Figure 28. Curvatures of the correlation function at the peak for different SNRs.

of the likelihood function depends on both the transmitted signal waveform $s(t)$ and the SNR, both of which are unknown. So, we need to use some curve fitting techniques around the peak to estimate the curvature of the likelihood function. We use the simplest form of fitting viz. quadratic fitting around the peak.

6.4 Results and Discussion

For the simulation, we used the setup shown in Figure 27. We used a Gaussian chirp as the unknown transmitted signal which is given by

$$s(t) = \exp\left(-\frac{1}{2}\sigma_F^2\left(t - \frac{T_s}{2}\right)^2\right) \sin(2\pi mt^2) \quad (52)$$

The signal is assumed to be approximately zero for $t < 0$ and for $t > T_s$. We set $T_s = 5\mu s$ and $\sigma_F = 2\pi \times 10^5$. The observation interval at each of the sensors was taken to be $T = 0.2$ ms. The unknown transmission time of the signal was set to $t_0 = 0.07$ ms. With this configuration the maximum TDOA is 0.0988 ms. The frequency spectrum of the Gaussian window is given by $|S(F)| = (\sqrt{2\pi}/\sigma_F) \exp(-2\pi^2 F^2/\sigma_F^2)$ [39], and the bandwidth of the chirp is $B_s = 1.5$ MHz. The rate of change of frequency for the linear chirp was chosen to be $m = \frac{B_s}{T_s} = 3 \times 10^8$. The correlation integrals are approximated using summations with $\delta t = 0.33$ ns which means the sampling frequency is $F_s = 300$ MHz. The reason for choosing such high sampling frequency is that at this frequency, the position quantization error due to sampling is in the order of $(c/F_s) = 10^{-3}$ km.

The regular two-step LS TDOA approach is implemented as follows. In the first step, the TDOA $\hat{\tau}'_i$ estimates were obtained by cross-correlating the signal at each sensor with the signal at sensor 0. Then a $1\text{km} \times 1\text{km}$ region around the true emitter location was split into 10×10 grid points and for each emitter location on

the grid point the TDOAs were computed using the formula

$$\tau'_i = \frac{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}}{c} - \frac{\sqrt{(x_T - x_0)^2 + (y_T - y_0)^2}}{c} \quad (53)$$

for $i = 1, 2, \dots, M - 1$. Next the LSE between the estimated TDOAs and the computed TDOAs was calculated as

$$\text{LSE} = \sum_{i=1}^{M-1} (\hat{\tau}'_i - \tau'_i)^2 \quad (54)$$

This LSE is a function of the emitter location (x_T, y_T) . The emitter location that minimized the LSE is the LS estimate of the emitter location. Since we know that the LSE is a 2-dimensional parabolic function of the emitter location near the minimum, we improved the accuracy by fitting a parabola through the 100 points at which the LSE was computed. The WLS estimator is also implemented in the same manner except the WLS error (WLSE) was calculated as

$$\text{WLSE} = \sum_{i=1}^{M-1} \alpha_i^2 (\hat{\tau}'_i - \tau'_i)^2 \quad (55)$$

where α_i^2 's are the weights. The emitter location that minimized the WLSE is the WLS estimate of the emitter location. In order to compute the weights, we chose a window of 30 points around the true peak of the correlation function. Using these 30 points we computed the coefficients (p_1, p_2, p_3) of a parabola $p_1 t^2 + p_2 t + p_3$ that fits these points in the least squares sense. Then the absolute value of the curvature of the parabola was taken as the weights. i.e,

$$\alpha = |2p_1| \quad (56)$$

We ran 1000 Monte Carlo simulations to compare the performance of the two estimators. The SNR was set at 0 dB at the reference sensor and at the other two sensors. At the fourth sensor, the SNR was varied from -30 dB to -20 dB. Figure 29 shows the comparison of the mean square error of the 1000 estimates for each of the two estimators for different SNR values of the fourth sensor.

Notice that the performance of the LS TDOA approach deteriorates significantly as the SNR goes below -28 dB. This is because at such low SNRs the TDOA estimator in (49) completely breaks down. When the noise level is low, the peak location of (51) is close to the true TDOA value but when the noise level is sufficiently high, the peak of (51) can occur anywhere in $(-T, T)$ with high probability. Figure 30 shows the histogram of the peak of the correlation function in (51) for the SNR values of 0 dB and -30 dB. Notice that at 0 dB the peak location lies very close to the true TDOA value of -0.0987 ms while at -30 dB, the peak of a realization appears at -0.06 ms also. When this outlier is used in the second step for position fixing, the location estimate is far from the true location. In the WLS TDOA approach, the effect of this outlier TDOA estimate is minimized by the weighting.

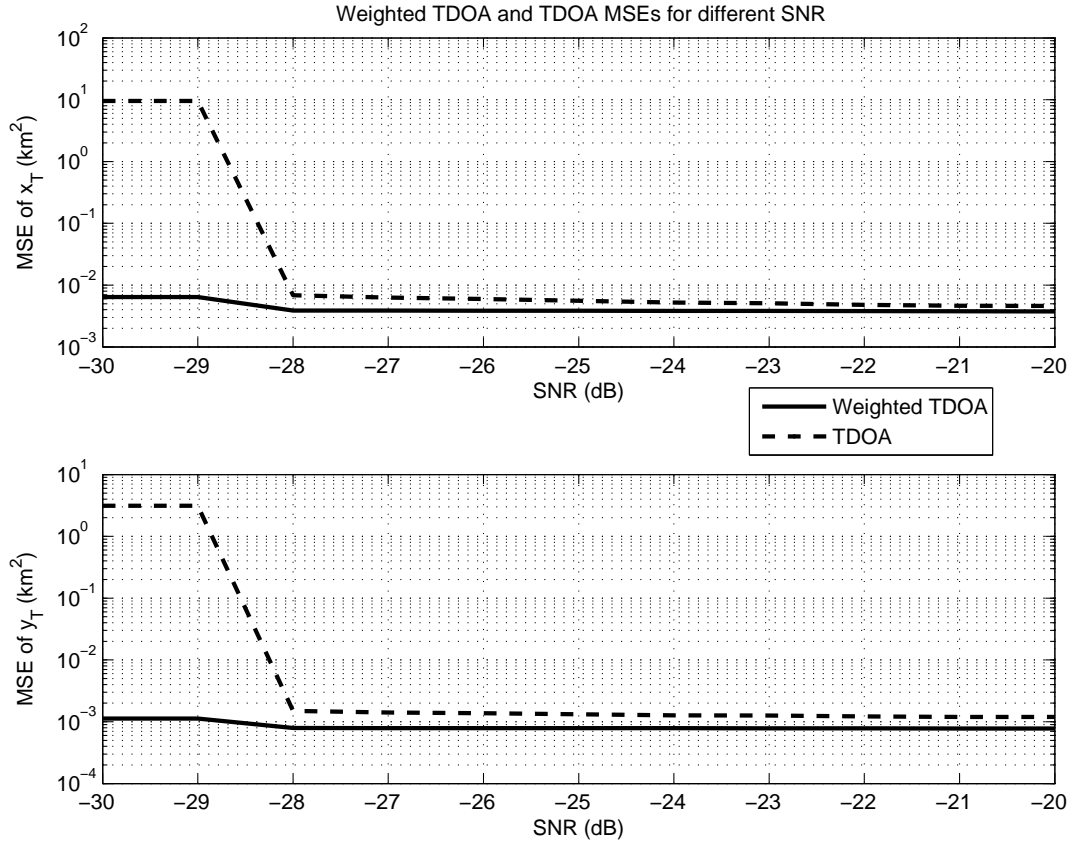


Figure 29. MSE of the 1000 estimates for each of the two estimators for different SNR values of the fourth sensor.

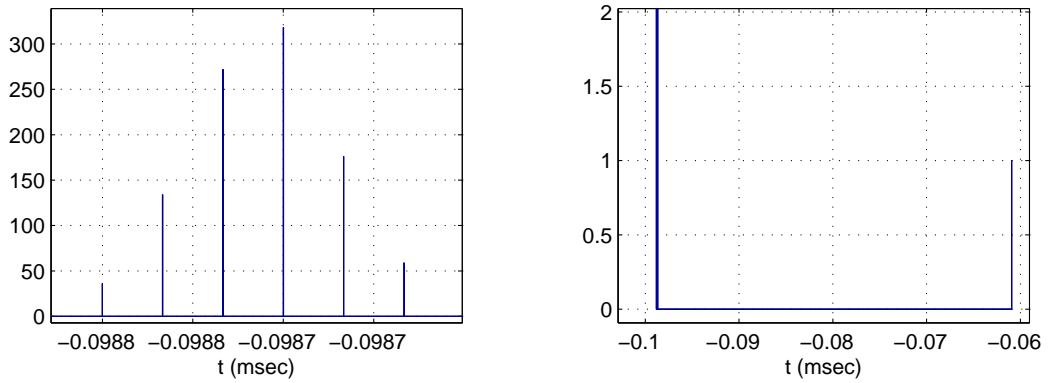


Figure 30. Histograms of the peak location: (a) SNR = 0 dB. (b) SNR = -30 dB.

6.5 Conclusion

We have proposed a weighted least squares position fixing technique for the multilateration problem. We showed that the TDOA estimator breaks down for

very low signal-to-noise ratios and when such TDOA estimates are used for position fixing, the location estimates are far from the true emitter location. Our weighting approach mitigates the effect of such outliers in the position fixing step. We proposed a simple technique for the computation of the weights from the received signal.

7 Overall Conclusion

We have addressed the problem of passively gathering information about an emitter of electronic signals using distributed sensors. First we proposed an asymptotically optimal technique for detection of the presence or absence of such signals in the data collected at the distributed sensors. This is a centralized detector unlike the commonly used decentralized decision fusion techniques. The derived detector is the generalized likelihood ratio test (GLRT) detector. We also derived simpler detectors from the GLRT by making various assumptions. Receiver operating characteristics (ROC) curves for currently used detectors are computed and compared to the ROC curves for the GLRT detector. After the presence of such signals is detected, the next step is to estimate the location of the emitter. For this we proposed the maximum likelihood estimator (MLE). The conventional approach for localization using multiple sensors is to first estimate the time difference of arrivals (TDOAs) of the signals, independently between pairs of sensors and then to find the location of the emitter using the intersection point of the hyperbolas defined by these TDOAs. This is referred to as the conventional TDOA technique and it has been shown in the literature that this two-step approach is suboptimal in comparison to what is called the direct position determination (DPD) approach. In the DPD approach, the intermediate step of estimating the TDOAs is bypassed and the location is estimated directly from the observations. In this paper we take the DPD approach instead of the conventional two-step approach. The DPD type localizers that have been proposed in the literature are based on certain assumptions on the transmitted signal such as narrowband or wideband, lowpass or bandpass etc. We make no such assumptions on the signal and this paper covers a wide variety of transmitted signals. In passive localization, it is common to not know the transmission time of the signal and more often the signal waveform itself is unknown. So, we have analyzed these two commonly occurring cases of (i) signal waveform unknown and (ii) signal waveform known with unknown transmission time. The localizers proposed in literature assumed discrete time but they have not addressed the quantization like effect on the location estimate due to sampling of the received signals. To avoid this quantization like effect, we have used a continuous time model. We have also derived the Fisher Information Matrix (FIM) which gives a deeper insight into the relations between various parameters in the model and their identifiability. We showed that the proposed MLE outperforms the conventional two-step localizers and also attains the Cramer Rao Lower Bound (CRLB) for high signal-to-noise ratios (SNR). Though the performance of the MLE attains the CRLB, there is still scope for improvement. This improvement comes from the geometry of the sensors. We showed that for a given SNR, the CRLB depends on the sensor geometry. Thus, the sensor geometry can be optimized in order to further reduce the CRLB and also the variance of the MLE. We have defined this problem of optimizing the sensor geometry and derived the

optimal sensor configurations for a few specific scenarios. The optimization of the sensor geometry for a general scenario is quite elusive and still open for research. In addition to efficient signal processing as in the MLE and optimizing the sensor geometries, the localization performance can further be enhanced by using information about the terrain. This is commonly referred to as the knowledge aided design (KAD). KAD is particularly useful for localization in urban environments where the distances are small and the knowledge of the terrain is very well known. We investigated the aspect of azimuth modulation induced by various objects in the terrain and the usefulness of this modulation for localization. In particular, we have shown that when an obstacle blocks the signal to one of the sensors, then there is increase in the over all information of the location of the emitter. All the concepts mentioned above are for improving the localization performance and do require altering/increasing the existing physical resources. For example the MLE requires high bandwidth links between all the sensors and the fusion center in order to transmit all the data collected at each sensor for simultaneous processing at the fusion center. This is in contrast to the conventional TDOA which only requires low bandwidth links to the fusion center because, here only the TDOA instead of the complete observation, is transmitted to the fusion center. So finally we propose an improvement to the conventional TDOA approach which does not require any additional physical resources but still significantly increases the localization performance particularly for SNRs at the break down range. The only additional piece of information that needs to be transmitted to the fusion center along with the TDOA is the curvature of the likelihood function. In the conventional TDOA approach the first step is to estimate the TDOAs and the second step, called the position fixing, is to estimate the emitter location as the intersection of the hyperboloids defined by these TDOAs. For the TDOA estimation the commonly used estimator is the time-delay that maximizes the cross-correlation function. This is the MLE of the TDOA (note that this is not the MLE of the emitter location which we have derived) and the cross-correlation function is the maximum likelihood function. Now, since the asymptotic variance of an MLE is equal to the negative of the expected value of the curvature of the likelihood function, we proposed a weighted least squares type position fixing technique where the weights can be computed from the curvature of the likelihood function. Hence, we have addressed the problem of passively gathering information about an emitter of electronic signals using distributed sensors and investigated various aspects that can increase this information.

REFERENCES

- [1] R. G. Wiley, *ELINT: The Interception and Analysis of Radar Signals*. Boston, MA: Artech, 2006.
- [2] P. Pace, *Detecting and Classifying Low Probability of Intercept Radar*. Norwood, MA: Artech House, 2004.
- [3] R. Tenney and N. Sandell, "Detection with distributed sensors," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-17, no. 4, pp. 501–510, July 1981.
- [4] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors i. fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, Jan 1997.
- [5] R. Srinivasan, "Distributed radar detection theory," *Communications, Radar and Signal Processing, IEE Proceedings F*, vol. 133, no. 1, pp. 55–60, Feb. 1986.
- [6] J.-F. Chamberland and V. Veeravalli, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1007–1015, Aug. 2004.
- [7] K. Liu and A. Sayeed, "Type-based decentralized detection in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 1899–1910, May 2007.
- [8] C. Xu and S. Kay, "On centralized composite detection with distributed sensors," in *Radar Conference, 2008. RADAR '08. IEEE*, 26-30 2008, pp. 1–6.
- [9] S. Stein, "Differential delay/Doppler ML estimation with unknown signals," *IEEE Transactions on Signal Processing*, vol. 41, no. 8, pp. 2717–2719, Aug 1993.
- [10] V. S. Chernyak, *Fundamentals of Multisite Radar Systems: Multistatic Radars and Multiradar Systems*. London, U.K.: Gordon and Breach, 1998.
- [11] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1998.
- [12] E. L. Lehmann, *Testing Statistical Hypothesis*. New York, NY: J. Wiley, 1959.

- [13] A. Weiss, "Direct position determination of narrowband radio transmitters," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04).*, vol. 2, May 2004, pp. ii – 249–50, ii–251–2 vol.2.
- [14] C. Holt, "Two-channel likelihood detectors for arbitrary linear channel distortion," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 35, no. 3, pp. 267 – 273, Mar 1987.
- [15] R. Altes, "Cross correlation and energy detection in multiarray processing," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 33, no. 3, pp. 493 – 504, Jun 1985.
- [16] K. R. Gabriel, "Least squares approximation of matrices by additive and multiplicative models," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 40, no. 2, pp. 186–196, 1978. [Online]. Available: <http://www.jstor.org/stable/2984755>
- [17] D. Torrieri, "Statistical theory of passive location systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-20, no. 2, pp. 183 –198, Mar. 1984.
- [18] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Transactions on Signal Processing*, vol. 42, no. 8, pp. 1905 –1915, Aug. 1994.
- [19] I. Mellen, G., M. Pachter, and J. Raquet, "Closed-form solution for determining emitter location using time difference of arrival measurements," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 3, pp. 1056 – 1058, Jul. 2003.
- [20] M. Gillette and H. Silverman, "A linear closed-form algorithm for source localization from time-differences of arrival," *Signal Processing Letters, IEEE*, vol. 15, pp. 1 – 4, 2008.
- [21] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 24, no. 4, pp. 320 – 327, Aug. 1976.
- [22] C. Knapp and G. Carter, "Time delay estimation in the presence of relative motion," in *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP '77.*, vol. 2, May. 1977, pp. 280 – 283.
- [23] A. Yeredor and E. Angel, "Joint TDOA and FDOA estimation: A conditional bound and its use for optimally weighted localization," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1612 –1623, April 2011.

- [24] W. R. Hahn, "Optimum signal processing for passive sonar range and bearing estimation," *The Journal of the Acoustical Society of America*, vol. 58, no. 1, pp. 201–207, 1975.
- [25] W. J. Bangs and P. M. Schultheiss, "Space-time processing for optimal parameter estimation," in *Signal Processing*, J. Griffiths, P.L. Stocklin, and C. V. Schooneveld, Eds. New York: Academic, 1973, pp. 577–590.
- [26] G. C. Carter, "Variance bounds for passively locating an acoustic source with a symmetric line array," *The Journal of the Acoustical Society of America*, vol. 62, no. 4, pp. 922–926, 1977.
- [27] B. Friedlander, "A passive localization algorithm and its accuracy analysis," *IEEE Journal of Oceanic Engineering*, vol. 12, no. 1, pp. 234–245, January 1987.
- [28] X. Hu and M. Fowler, "Sensor selection for multiple sensor emitter location systems," in *Aerospace Conference, 2008 IEEE*, Mar. 2008, pp. 1–10.
- [29] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 35, no. 8, pp. 1223–1225, Aug 1987.
- [30] E. Xu, Z. Ding, and S. Dasgupta, "Source localization in wireless sensor networks from signal time-of-arrival measurements," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2887–2897, June 2011.
- [31] J.-Y. Do, M. Rabinowitz, and P. Enge, "Performance of TOA and TDOA in a non-homogeneous transmitter network combining GPS and terrestrial signals," in *Proceedings of the ION National Technical Meeting, Monterey, CA*, Jan. 2006, pp. 642–649.
- [32] A. Weiss, "Direct position determination of narrowband radio frequency transmitters," *Signal Processing Letters, IEEE*, vol. 11, no. 5, pp. 513–516, May 2004.
- [33] A. Amar and A. Weiss, "Direct position determination of multiple radio signals," in *Proceedings on IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '04), 2004.*, vol. 2, 17–21 2004, pp. ii–81–4 vol.2.
- [34] A. Amar and A. Weiss, "Localization of narrowband radio emitters based on Doppler frequency shifts," *IEEE Transactions on Signal Processing*, vol. 56, no. 11, pp. 5500–5508, Nov. 2008.

- [35] A. Weiss and A. Amar, "Direct geolocation of stationary wideband radio signal based on time delays and Doppler shifts," in *IEEE/SP 15th Workshop on Statistical Signal Processing, 2009. SSP '09*, Aug. 2009, pp. 101–104.
- [36] H. Lee, "Accuracy limitations of hyperbolic multilateration systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-11, no. 1, pp. 16–29, jan. 1975.
- [37] A. Bishop, B. Fidan, B. Anderson, P. Pathirana, and K. Dogancay, "Optimality analysis of sensor-target geometries in passive localization: Part 2 - time-of-arrival based localization," in *Intelligent Sensors, Sensor Networks and Information, 2007. ISSNIP 2007. 3rd International Conference on*, dec. 2007, pp. 13–18.
- [38] A. D. Whalen, *Detection of Signals in Noise*. Academic Press, New York, 1971.
- [39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1993.
- [40] P. Stoica and T. Marzetta, "Parameter estimation problems with singular information matrices," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 87–90, Jan. 2001.
- [41] D. Musicki, R. Kaune, and W. Koch, "Mobile emitter geolocation and tracking using TDOA and FDOA measurements," *IEEE Transactions on Signal Processing*, vol. 58, no. 3, pp. 1863–1874, March 2010.
- [42] W. Hahn and S. Tretter, "Optimum processing for delay-vector estimation in passive signal arrays," *IEEE Transactions on Information Theory*, vol. 19, no. 5, pp. 608–614, sep 1973.
- [43] K. Ho and Y. Chan, "Solution and performance analysis of geolocation by tdoa," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1311–1322, oct 1993.
- [44] A. Mathias, M. Leonardi, and G. Galati, "An efficient multilateration algorithm," in *Digital Communications - Enhanced Surveillance of Aircraft and Vehicles, 2008. TIWDC/ESAV 2008. Tyrrhenian International Workshop on*, sept. 2008, pp. 1–6.

A Derivation of the GLRT Detector

The GLRT decides \mathcal{H}_1 if

$$L_G(\tilde{\mathbf{r}}) = \frac{p(\tilde{\mathbf{r}}; \hat{\tilde{\mathbf{s}}}, \hat{\tilde{\mathbf{A}}}, \hat{\mathbf{n}}, \hat{\mathbf{k}}, \mathcal{H}_1)}{p(\tilde{\mathbf{r}}; \mathcal{H}_0)} > \gamma$$

or equivalently if the test statistic

$$\ln L_G(\tilde{\mathbf{r}}) > \ln \gamma = \gamma'$$

Here $\hat{\tilde{\mathbf{s}}}$, $\hat{\tilde{\mathbf{A}}}$, $\hat{\mathbf{n}}$ and $\hat{\mathbf{k}}$ are the MLEs of $\tilde{\mathbf{s}}$, $\tilde{\mathbf{A}}$, \mathbf{n} and \mathbf{k} respectively assuming \mathcal{H}_1 is true. Since $\tilde{\mathbf{w}} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{MN})$ we have

$$L_G(\tilde{\mathbf{r}}) = \frac{\frac{1}{\pi^{NM}} \exp \left(- \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \hat{\tilde{\mathbf{s}}} \right)^H \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \hat{\tilde{\mathbf{s}}} \right) \right)}{\frac{1}{\pi^{NM}} \exp \left(- \tilde{\mathbf{r}}^H \tilde{\mathbf{r}} \right)}$$

Taking logarithms we have

$$\begin{aligned} \ln L_G(\tilde{\mathbf{r}}) &= - \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\hat{\tilde{\mathbf{A}}}, \hat{\mathbf{n}}, \hat{\mathbf{k}}) \hat{\tilde{\mathbf{s}}} \right)^H \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\hat{\tilde{\mathbf{A}}}, \hat{\mathbf{n}}, \hat{\mathbf{k}}) \hat{\tilde{\mathbf{s}}} \right) + \tilde{\mathbf{r}}^H \tilde{\mathbf{r}} \\ &= 2\text{Re} \left[\tilde{\mathbf{r}}^H \tilde{\mathbf{H}}(\hat{\tilde{\mathbf{A}}}, \hat{\mathbf{n}}, \hat{\mathbf{k}}) \hat{\tilde{\mathbf{s}}} \right] - \left(\tilde{\mathbf{H}}(\hat{\tilde{\mathbf{A}}}, \hat{\mathbf{n}}, \hat{\mathbf{k}}) \hat{\tilde{\mathbf{s}}} \right)^H \left(\tilde{\mathbf{H}}(\hat{\tilde{\mathbf{A}}}, \hat{\mathbf{n}}, \hat{\mathbf{k}}) \hat{\tilde{\mathbf{s}}} \right) \end{aligned}$$

The MLEs $\hat{\tilde{\mathbf{s}}}$, $\hat{\tilde{\mathbf{A}}}$, $\hat{\mathbf{n}}$ and $\hat{\mathbf{k}}$ maximize $p(\tilde{\mathbf{r}}; \tilde{\mathbf{s}}, \tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}; \mathcal{H}_1)$. But,

$$p(\tilde{\mathbf{r}}; \tilde{\mathbf{s}}, \tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}, \mathcal{H}_1) = \frac{1}{\pi^{NM}} \exp \left[- \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{s}} \right)^H \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{s}} \right) \right]$$

Maximizing $p(\tilde{\mathbf{r}}; \tilde{\mathbf{s}}, \tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}, \mathcal{H}_1)$ over $\tilde{\mathbf{s}}$, $\tilde{\mathbf{A}}$, \mathbf{n} and \mathbf{k} is equivalent to minimizing the exponent which is

$$J = \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{s}} \right)^H \left(\tilde{\mathbf{r}} - \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{s}} \right)$$

From [39] we have the MLE of $\tilde{\mathbf{s}}$ as

$$\hat{\tilde{\mathbf{s}}} = \left[\tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \right]^{-1} \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{r}}$$

Putting this back in the test statistic we have

$$\ln L_G(\tilde{\mathbf{r}}) = \max_{\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}} \left[\tilde{\mathbf{r}}^H \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \left[\tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \right]^{-1} \tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{r}} \right]$$

Notice that

$$\tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) = (\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}) \mathbf{I}_N$$

and

$$\tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{r}} = \left(\sum_{i=0}^{M-1} \tilde{A}_i^* (\tilde{\mathbf{W}}^{k_i})^H (\mathbf{P}^{n_i})^H \tilde{\mathbf{r}}_i \right)$$

so that we have

$$\begin{aligned} \ln L_G(\tilde{\mathbf{r}}) &= \max_{\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}} \frac{\tilde{\mathbf{r}}^H \tilde{\mathbf{H}}(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{r}}}{(\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})} \\ &= \max_{\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}} \frac{\left(\tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{r}} \right)^H \left(\tilde{\mathbf{H}}^H(\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}) \tilde{\mathbf{r}} \right)}{(\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})} \\ &= \max_{\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}} \frac{\left(\sum_{i=0}^{M-1} \tilde{A}_i^* (\tilde{\mathbf{W}}^{k_i})^H (\mathbf{P}^{n_i})^H \tilde{\mathbf{r}}_i \right)^H \left(\sum_{i=0}^{M-1} \tilde{A}_i^* (\tilde{\mathbf{W}}^{k_i})^H (\mathbf{P}^{n_i})^H \tilde{\mathbf{r}}_i \right)}{(\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})} \\ &= \max_{\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}} \frac{\left(\sum_{i=0}^{M-1} \tilde{A}_i \tilde{\mathbf{r}}_i^H (\mathbf{P}^{n_i}) (\tilde{\mathbf{W}}^{k_i}) \right) \left(\sum_{i=0}^{M-1} \tilde{A}_i^* (\tilde{\mathbf{W}}^{k_i})^H (\mathbf{P}^{n_i})^H \tilde{\mathbf{r}}_i \right)}{(\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})} \\ &= \max_{\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}} \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \tilde{A}_i \tilde{\mathbf{r}}_i^H (\mathbf{P}^{n_i}) (\tilde{\mathbf{W}}^{k_i}) (\tilde{\mathbf{W}}^{k_j})^H (\mathbf{P}^{n_j})^H \tilde{\mathbf{r}}_j \tilde{A}_j^*}{(\tilde{\mathbf{A}}^H \tilde{\mathbf{A}})} \end{aligned}$$

If we let $\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k})$ be the cross ambiguity matrix defined as

$$\begin{aligned} \tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k}) &= \begin{bmatrix} \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} \\ \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} \\ \vdots \\ \tilde{\mathbf{r}}_{M-1}^H \mathbf{P}^{n_{M-1}} \tilde{\mathbf{W}}^{k_{M-1}} \end{bmatrix} \cdot \begin{bmatrix} \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} \\ \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} \\ \vdots \\ \tilde{\mathbf{r}}_{M-1}^H \mathbf{P}^{n_{M-1}} \tilde{\mathbf{W}}^{k_{M-1}} \end{bmatrix}^H \\ &= \begin{bmatrix} \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 & \dots & \tilde{\mathbf{r}}_0^H \mathbf{P}^{n_0} \tilde{\mathbf{W}}^{k_0} (\tilde{\mathbf{W}}^{k_{M-1}})^H (\mathbf{P}^{n_{M-1}})^H \tilde{\mathbf{r}}_{M-1} \\ \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 & \dots & \tilde{\mathbf{r}}_1^H \mathbf{P}^{n_1} \tilde{\mathbf{W}}^{k_1} (\tilde{\mathbf{W}}^{k_{M-1}})^H (\mathbf{P}^{n_{M-1}})^H \tilde{\mathbf{r}}_{M-1} \\ \vdots & \vdots & \vdots \\ \tilde{\mathbf{r}}_{M-1}^H \mathbf{P}^{n_{M-1}} \tilde{\mathbf{W}}^{k_{M-1}} (\tilde{\mathbf{W}}^{k_0})^H (\mathbf{P}^{n_0})^H \tilde{\mathbf{r}}_0 & \dots & \tilde{\mathbf{r}}_{M-1}^H \mathbf{P}^{n_{M-1}} \tilde{\mathbf{W}}^{k_{M-1}} (\tilde{\mathbf{W}}^{k_{M-1}})^H (\mathbf{P}^{n_{M-1}})^H \tilde{\mathbf{r}}_{M-1} \end{bmatrix}_{M \times M} \end{aligned}$$

then we have a Hermitian form and so

$$\ln L_G(\tilde{\mathbf{r}}) = \max_{\tilde{\mathbf{A}}, \mathbf{n}, \mathbf{k}} \frac{\tilde{\mathbf{A}}^H \tilde{\mathbf{B}}^*(\mathbf{n}, \mathbf{k}) \tilde{\mathbf{A}}}{\tilde{\mathbf{A}}^H \tilde{\mathbf{A}}} = \max_{\mathbf{n}, \mathbf{k}} \lambda_{\max}(\tilde{\mathbf{B}}^*(\mathbf{n}, \mathbf{k}))$$

where λ_{\max} is the maximum eigenvalue of $\tilde{\mathbf{B}}^*(\mathbf{n}, \mathbf{k})$. The MLE of $\tilde{\mathbf{A}}$ is the eigenvector corresponding to $\lambda_{\max}(\tilde{\mathbf{B}}^*(\mathbf{n}, \mathbf{k}))$. Therefore, the GLRT decides \mathcal{H}_1 if

$$\max_{\mathbf{n}, \mathbf{k}} \lambda_{\max}(\tilde{\mathbf{B}}^*(\mathbf{n}, \mathbf{k})) > \gamma'$$

The cross ambiguity matrix $\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k})$ is Hermitian with real and positive eigenvalues. So, we have $\lambda_{\max}(\tilde{\mathbf{B}}^*(\mathbf{n}, \mathbf{k})) = \lambda_{\max}(\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k}))$ and so the GLRT decides \mathcal{H}_1 if

$$\max_{\mathbf{n}, \mathbf{k}} \lambda_{\max}(\tilde{\mathbf{B}}(\mathbf{n}, \mathbf{k})) > \gamma'$$

B Localization

B.1 CRLB

We will derive the CRLB for the emitter location estimate. First, we will show that the FIM for the model used for unknown signal with unknown transmission time case is singular. We will then use a transformation of the parameters in the model and derive the CRLB. Let $\boldsymbol{\tau} = [\tau_0 \ \tau_1 \ \cdots \ \tau_{M-1}]^T$, $\mathbf{A} = [A_0 \ A_1 \ \cdots \ A_{M-1}]^T$ and $\boldsymbol{\phi} = [a_0 \ a_1 \ \cdots \ a_{N-1} \ b_1 \ b_2 \ \cdots \ b_{N-1}]^T$, where

$$a_0 = \frac{\sqrt{2}}{T} \int_0^T s(t) dt, \quad a_n = \frac{2}{T} \int_0^T s(t) \cos 2\pi n F_0 t dt, \quad b_n = \frac{2}{T} \int_0^T s(t) \sin 2\pi n F_0 t dt$$

Let $\boldsymbol{\theta} = [\boldsymbol{\tau}^T \ \mathbf{A}^T \ \boldsymbol{\phi}^T]^T$. The TOAs τ_i are a function of the emitter location (x_T, y_T) and the signal transmission time t_0 .

$$\tau_i = \frac{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}}{c} + t_0$$

where c is the propagation speed of the signal. If $l(\boldsymbol{\theta})$ is the log-likelihood function, then the FIM is given by

$$\mathcal{I}_\theta = \begin{bmatrix} -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \mathbf{A}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\phi}^T} \right\} \\ -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mathbf{A} \partial \boldsymbol{\tau}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mathbf{A} \partial \mathbf{A}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mathbf{A} \partial \boldsymbol{\phi}^T} \right\} \\ -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\tau}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \mathbf{A}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} \right\} \end{bmatrix} \quad (\text{B.1})$$

B.1.1 Signal unknown with unknown transmission time

From (23), we have the log-likelihood function as

$$l(\boldsymbol{\theta}) = -\frac{1}{N_0} \int_0^T \sum_{m=0}^{M-1} (x_m(t) - A_m \mathbf{h}^T(t - \tau_m) \boldsymbol{\phi})^2 dt \quad (\text{B.2})$$

where $\mathbf{h}(t)$ is as defined in Appendix B.3. Partial differentiation with respect to (w.r.t) τ_i gives,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \tau_i} = -\frac{1}{N_0} \int_0^T 2 (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi}) \left(-A_i \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \boldsymbol{\phi} \right) dt \quad (\text{B.3})$$

and w.r.t A_i gives,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial A_i} = -\frac{1}{N_0} \int_0^T 2 (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi}) (-\mathbf{h}^T(t - \tau_i) \boldsymbol{\phi}) dt \quad (\text{B.4})$$

for $i = 0, 1, \dots, M-1$. Partial differentiation w.r.t ϕ gives,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial \phi} = -\frac{1}{N_0} \int_0^T \sum_{m=0}^{M-1} 2(x_m(t) - A_m \mathbf{h}^T(t - \tau_m) \phi) (-A_m \mathbf{h}^T(t - \tau_m)) dt \quad (\text{B.5})$$

Next we evaluate the second derivatives. Partial differentiation of (B.3) w.r.t τ_i gives,

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i^2} &= -\frac{2}{N_0} \left[\int_0^T (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \phi) \left(-A_i \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \phi \right) dt \right. \\ &\quad \left. + \int_0^T \left(-A_i \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \phi \right)^2 dt \right]. \end{aligned}$$

Taking the negative of the expected value of both sides and using (B.27) gives,

$$\begin{aligned} -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i^2} \right\} &= \frac{2}{N_0} \int_0^T \left(A_i \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \phi \right)^2 dt \\ &= \frac{A_i^2}{(N_0/2)} \phi^T \left[\int_0^T \frac{\partial \mathbf{h}(t - \tau_i)}{\partial \tau_i} \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \right] \phi \\ &= \frac{(T/2)(2\pi F_0)^2}{(N_0/2)} (\phi^T \mathbf{L} \mathbf{L}^T \phi) A_i^2 \end{aligned}$$

where the $2N-1 \times 2N-1$ matrix \mathbf{L} is as defined in Appendix B.3. Since $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i \partial \tau_j} = 0$ for $i \neq j$, we have

$$\boxed{-E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}^T} \right\} = \frac{(T/2)(2\pi F_0)^2}{(N_0/2)} \phi^T \mathbf{L} \mathbf{L}^T \phi (\text{diag}(\mathbf{A}))^2} \quad (\text{B.6})$$

where $\text{diag}(\mathbf{A})$ is an $M \times M$ diagonal matrix with i th diagonal element as A_i . Partial differentiation of (B.4) w.r.t τ_i gives,

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i \partial A_i} &= -\frac{2}{N_0} \left[\int_0^T (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \phi) \left(-\frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \phi \right) dt + \right. \\ &\quad \left. \int_0^T \left(-A_i \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \phi \right) (-\mathbf{h}^T(t - \tau_i) \phi) dt \right] \end{aligned}$$

Taking the negative of the expected value of both sides and using (B.28) gives,

$$\begin{aligned} -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i \partial A_i} \right\} &= \frac{A_i}{(N_0/2)} \phi^T \left[\int_0^T \frac{\partial \mathbf{h}(t - \tau_i)}{\partial \tau_i} \mathbf{h}^T(t - \tau_i) dt \right] \phi \\ &= \frac{(T/2)(2\pi F_0)}{(N_0/2)} (\phi^T \mathbf{L} \phi) A_i \end{aligned}$$

Since $\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i \partial A_j} = 0$ for $i \neq j$, we have

$$\boxed{-\mathbb{E} \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \mathbf{A}^T} \right\} = \frac{(T/2)(2\pi F_0)}{(N_0/2)} (\boldsymbol{\phi}^T \mathbf{L} \boldsymbol{\phi}) (\text{diag}(\mathbf{A}))} \quad (\text{B.7})$$

Partial differentiation of (B.5) w.r.t τ_i gives,

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i \partial \boldsymbol{\phi}^T} = -\frac{2}{N_0} \left[\int_0^T (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi}) \left(-A_i \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \right) dt + \int_0^T \left(-A_i \frac{\partial \mathbf{h}^T(t - \tau_i)}{\partial \tau_i} \boldsymbol{\phi} \right) (-A_i \mathbf{h}^T(t - \tau_i)) dt \right]$$

Taking the negative of the expected value of both sides and using (B.28) gives,

$$\begin{aligned} -\mathbb{E} \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \tau_i \partial \boldsymbol{\phi}^T} \right\} &= \frac{A_i^2}{(N_0/2)} \boldsymbol{\phi}^T \left[\int_0^T \frac{\partial \mathbf{h}(t - \tau_i)}{\partial \tau_i} \mathbf{h}^T(t - \tau_i) dt \right] \\ &= \frac{(T/2)(2\pi F_0) A_i^2}{(N_0/2)} \boldsymbol{\phi}^T \mathbf{L} \end{aligned}$$

So, we have

$$\boxed{-\mathbb{E} \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\phi}^T} \right\} = \frac{(T/2)(2\pi F_0)}{(N_0/2)} (\mathbf{A} \odot \mathbf{A}) \boldsymbol{\phi}^T \mathbf{L}} \quad (\text{B.8})$$

Partial differentiation of (B.4) w.r.t A_i and using (B.26) gives,

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial A_i^2} &= -\frac{1}{(N_0/2)} \int_0^T (\mathbf{h}^T(t - \tau_i) \boldsymbol{\phi})^2 \\ &= -\frac{1}{(N_0/2)} \boldsymbol{\phi}^T \left[\int_0^T \mathbf{h}(t - \tau_i) \mathbf{h}^T(t - \tau_i) \right] \boldsymbol{\phi} \\ &= -\frac{(T/2)}{(N_0/2)} \boldsymbol{\phi}^T \boldsymbol{\phi} \end{aligned}$$

Taking the negative of the expected value on both sides gives,

$$\boxed{-\mathbb{E} \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mathbf{A} \partial \mathbf{A}^T} \right\} = \frac{(T/2) \boldsymbol{\phi}^T \boldsymbol{\phi}}{(N_0/2)} \mathbf{I}_M} \quad (\text{B.9})$$

Partial differentiation of (B.5) w.r.t A_i gives,

$$\frac{\partial^2 l(\boldsymbol{\theta})}{\partial A_i \partial \boldsymbol{\phi}^T} = -\frac{2}{N_0} \left[\int_0^T (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi}) (-\mathbf{h}^T(t - \tau_i)) dt + \int_0^T (-\mathbf{h}^T(t - \tau_i) \boldsymbol{\phi}) (-A_i \mathbf{h}^T(t - \tau_i)) dt \right]$$

Taking the negative of the expected value of both sides and using (B.26) gives,

$$\begin{aligned} -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial A_i \partial \boldsymbol{\phi}^T} \right\} &= \frac{2A_i}{N_0} \int_0^T \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi} \mathbf{h}^T(t - \tau_i) dt \\ &= \frac{2A_i}{N_0} \boldsymbol{\phi}^T \int_0^T \mathbf{h}(t - \tau_i) \mathbf{h}^T(t - \tau_i) dt \\ &= \frac{(T/2)}{(N_0/2)} A_i \boldsymbol{\phi}^T \end{aligned}$$

and so

$$\boxed{-E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mathbf{A} \partial \boldsymbol{\phi}^T} \right\} = \frac{(T/2)}{(N_0/2)} \mathbf{A} \boldsymbol{\phi}^T} \quad (\text{B.10})$$

Partial differentiation of (B.5) w.r.t $\boldsymbol{\phi}$ and using (B.26) gives,

$$\begin{aligned} \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} &= -\frac{2}{N_0} \int_0^T \sum_{i=0}^{M-1} (-A_i \mathbf{h}^T(t - \tau_i)) (-A_i \mathbf{h}^T(t - \tau_i)) dt \\ &= -\sum_{i=0}^{M-1} \frac{A_i^2}{(N_0/2)} \int_0^T \mathbf{h}(t - \tau_i) \mathbf{h}^T(t - \tau_i) dt \\ &= -\sum_{i=0}^{M-1} \frac{(T/2)}{(N_0/2)} \mathbf{I}_{(2N-1)} A_i^2 \end{aligned}$$

and so

$$\boxed{-E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\phi} \partial \boldsymbol{\phi}^T} \right\} = \frac{(T/2)}{(N_0/2)} \mathbf{I}_{(2N-1)} \sum_{i=0}^{M-1} A_i^2 = \frac{(T/2) \mathbf{A}^T \mathbf{A}}{(N_0/2)} \mathbf{I}_{(2N-1)}} \quad (\text{B.11})$$

Putting (B.6), (B.7), (B.8), (B.9), (B.10), (B.11) back in (B.1), we have

$$\mathcal{I}_\theta = \frac{(T/2)}{(N_0/2)} \begin{bmatrix} (2\pi F_0)^2 \boldsymbol{\phi}^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi} (\text{diag}(\mathbf{A}))^2 & (2\pi F_0)(\boldsymbol{\phi}^T \mathbf{L} \boldsymbol{\phi}) (\text{diag}(\mathbf{A})) & (2\pi F_0)(\mathbf{A} \odot \mathbf{A}) \boldsymbol{\phi}^T \mathbf{L} \\ (2\pi F_0)(\boldsymbol{\phi}^T \mathbf{L}^T \boldsymbol{\phi}) (\text{diag}(\mathbf{A})) & (\boldsymbol{\phi}^T \boldsymbol{\phi}) \mathbf{I}_M & \mathbf{A} \boldsymbol{\phi}^T \\ (2\pi F_0) \mathbf{L}^T \boldsymbol{\phi} (\mathbf{A} \odot \mathbf{A})^T & \boldsymbol{\phi} \mathbf{A}^T & (\mathbf{A}^T \mathbf{A}) \mathbf{I}_{(2N-1)} \end{bmatrix} \quad (\text{B.12})$$

The CRLB matrix for the unknown parameter vector $\boldsymbol{\theta}$ is the inverse of the matrix \mathcal{I}_θ . But in Appendix B.4 it is shown that the null space of \mathcal{I}_θ is not empty and so it is not invertible. This is because the log-likelihood function is not uniquely defined by the model in (23). To eliminate the over parameterization we use the following transformations.

$$\begin{aligned}\boldsymbol{\tau}' &= [(\tau_1 - \tau_0) \quad (\tau_2 - \tau_0) \quad \cdots \quad (\tau_{M-1} - \tau_0)]^T \\ \mathbf{A}' &= (1/A_0)[A_1 \quad \cdots \quad A_{M-1}]^T \\ \boldsymbol{\phi}' &= A_0 \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}(\mathbf{h}(-\tau_0))\boldsymbol{\phi}\end{aligned}\quad (\text{B.13})$$

Let $\boldsymbol{\theta}' = [\boldsymbol{\tau}'^T \quad \mathbf{A}'^T \quad \boldsymbol{\phi}'^T]^T$. This is a function of $\boldsymbol{\theta}$. Let $\mathbf{H} = \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\theta}}\right)$ be the Jacobian. If \mathbf{H} has row vectors that are linear combinations of those eigenvectors of \mathcal{I}_θ that have nonzero eigenvalues then the CRLB of $\boldsymbol{\theta}'$ is given by $\mathbf{H}\mathcal{I}_\theta^\dagger \mathbf{H}^T$ [40]. The \dagger is used to represent the generalized inverse. This condition is verified in Appendix B.4. Therefore,

$$\mathcal{I}_{\theta'}^{-1} = \mathbf{H}\mathcal{I}_\theta^\dagger \mathbf{H}^T \quad (\text{B.14})$$

Alternately, the log-likelihood function for this model with the transformed parameters is given by

$$l(\boldsymbol{\theta}') = -\frac{1}{N_0} \int_0^T (r_0(t) - \mathbf{h}^T(t)\boldsymbol{\phi}')^2 dt - \frac{1}{N_0} \int_0^T \sum_{i=1}^{M-1} (r_i(t) - A_i' \mathbf{h}^T(t - \tau_i')\boldsymbol{\phi}')^2 dt \quad (\text{B.15})$$

So, computing the derivatives to find the FIM as done previously yields the FIM for the transformed parameter vector as

$$\mathcal{I}_{\theta'} = \frac{T/2}{(N_0/2)} \begin{bmatrix} (2\pi F_0)^2 \boldsymbol{\phi}'^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi}' (\text{diag}(\mathbf{A}'))^2 & (2\pi F_0)(\boldsymbol{\phi}'^T \mathbf{L} \boldsymbol{\phi}') (\text{diag}(\mathbf{A}')) & (2\pi F_0)(\mathbf{A}' \odot \mathbf{A}') \boldsymbol{\phi}'^T \mathbf{L} \\ (2\pi F_0)(\boldsymbol{\phi}'^T \mathbf{L} \boldsymbol{\phi}') (\text{diag}(\mathbf{A}')) & (\boldsymbol{\phi}'^T \boldsymbol{\phi}') \mathbf{I}_{M-1} & \mathbf{A}' \boldsymbol{\phi}'^T \\ (2\pi F_0) \mathbf{L}^T \boldsymbol{\phi}' (\mathbf{A}' \odot \mathbf{A}')^T & \boldsymbol{\phi}' \mathbf{A}'^T & (1 + \mathbf{A}'^T \mathbf{A}') \mathbf{I}_{(2N-1)} \end{bmatrix} \quad (\text{B.16})$$

We have verified numerically that (B.14) is equivalent to (B.16). The elements of the TDOA vector $\boldsymbol{\tau}'$ are given by

$$\tau_i' = (\tau_i - \tau_0) = \frac{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}}{c} - \frac{\sqrt{(x_T - x_0)^2 + (y_T - y_0)^2}}{c}$$

so that the new parameter vector $\boldsymbol{\tau}'$ is a function of only the emitter location (x_T, y_T) . So, if we let $\boldsymbol{\eta}' = [x_T \ y_T]^T$ and $\boldsymbol{\alpha}' = [\boldsymbol{\eta}'^T \quad \mathbf{A}'^T \quad \boldsymbol{\phi}'^T]^T$ we have

$$\mathcal{I}_{\alpha'} = \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T}\right)^T \mathcal{I}_{\theta'} \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T}\right) = \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T}\right)^T (\mathbf{H}\mathcal{I}_\theta^\dagger \mathbf{H}^T)^{-1} \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T}\right) \quad (\text{B.17})$$

The Jacobian is given by the $(2M + 2N - 3, M + 2N)$ matrix

$$\begin{aligned} \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\alpha}'^T} \right) &= \begin{bmatrix} \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) & \left(\frac{\partial \boldsymbol{\tau}'}{\partial \mathbf{A}'^T} \right) & \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\phi}'^T} \right) \\ \left(\frac{\partial \mathbf{A}'}{\partial \boldsymbol{\eta}'^T} \right) & \left(\frac{\partial \mathbf{A}'}{\partial \mathbf{A}'^T} \right) & \left(\frac{\partial \mathbf{A}'}{\partial \boldsymbol{\phi}'^T} \right) \\ \left(\frac{\partial \boldsymbol{\phi}'}{\partial \boldsymbol{\eta}'^T} \right) & \left(\frac{\partial \boldsymbol{\phi}'}{\partial \mathbf{A}'^T} \right) & \left(\frac{\partial \boldsymbol{\phi}'}{\partial \boldsymbol{\phi}'^T} \right) \end{bmatrix} \\ &= \begin{bmatrix} \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) & \mathbf{0}_{(M,M)} & \mathbf{0}_{(M,2N-1)} \\ \mathbf{0}_{(M,2)} & \mathbf{I}_{M-1} & \mathbf{0}_{(M,2N-1)} \\ \mathbf{0}_{(2N-1,2)} & \mathbf{0}_{(2N-1,M)} & \mathbf{I}_{(2N-1)} \end{bmatrix} \end{aligned} \quad (\text{B.18})$$

where the $(M - 1) \times 2$ matrix $\left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) =$

$$(1/c) \begin{bmatrix} \frac{(x_T - x_1)}{d_1} - \frac{(x_T - x_0)}{d_0} & \frac{(y_T - y_1)}{d_1} - \frac{(y_T - y_0)}{d_0} \\ \frac{(x_T - x_2)}{d_2} - \frac{(x_T - x_0)}{d_0} & \frac{(y_T - y_2)}{d_2} - \frac{(y_T - y_0)}{d_0} \\ \vdots & \vdots \\ \frac{(x_T - x_{M-1})}{d_{M-1}} - \frac{(x_T - x_0)}{d_0} & \frac{(y_T - y_{M-1})}{d_{M-1}} - \frac{(y_T - y_0)}{d_0} \end{bmatrix} \quad (\text{B.19})$$

where d_i , $i = 0, \dots, M - 1$ is the distance between the sensor i and the emitter.

B.1.2 Signal known with unknown transmission time

From (34) we have the log-likelihood function as

$$l(\boldsymbol{\zeta}) = -\frac{1}{N_0} \int_0^T \sum_{i=0}^{M-1} (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi})^2 dt \quad (\text{B.20})$$

Here the $2M \times 1$ unknown parameter vector is $\boldsymbol{\zeta} = [\boldsymbol{\tau}^T \mathbf{A}^T]^T$. So, the FIM is given by

$$\mathcal{I}_{\boldsymbol{\zeta}} = \begin{bmatrix} -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \boldsymbol{\tau}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\tau} \partial \mathbf{A}^T} \right\} \\ -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mathbf{A} \partial \boldsymbol{\tau}^T} \right\} & -E \left\{ \frac{\partial^2 l(\boldsymbol{\theta})}{\partial \mathbf{A} \partial \mathbf{A}^T} \right\} \end{bmatrix} \quad (\text{B.21})$$

Using (B.6), (B.7) and (B.9), we have

$$\mathcal{I}_\zeta = \frac{(T/2)}{(N_0/2)} \begin{bmatrix} (2\pi F_0)^2 \boldsymbol{\phi}^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi} (\text{diag}(\mathbf{A}))^2 & (2\pi F_0)(\boldsymbol{\phi}^T \mathbf{L} \boldsymbol{\phi})(\text{diag}(\mathbf{A})) \\ (2\pi F_0)(\boldsymbol{\phi}^T \mathbf{L}^T \boldsymbol{\phi})(\text{diag}(\mathbf{A})) & (\boldsymbol{\phi}^T \boldsymbol{\phi}) \mathbf{I}_M \end{bmatrix} \quad (\text{B.22})$$

B.2 Maximum Likelihood Estimator

Here we will derive the MLE for the two cases of signal unknown with unknown transmission time and signal known with unknown transmission time.

B.2.1 Signal unknown with unknown transmission time

The log-likelihood function with the transformed parameters is given by

$$l(\boldsymbol{\theta}') = -\frac{1}{N_0} \int_0^T (r_0(t) - \mathbf{h}^T(t) \boldsymbol{\phi}')^2 dt - \frac{1}{N_0} \int_0^T \sum_{i=1}^{M-1} (r_i(t) - A'_i \mathbf{h}^T(t - \tau'_i) \boldsymbol{\phi}')^2 dt \quad (\text{B.23})$$

Partial differentiation w.r.t $\boldsymbol{\phi}'$ gives,

$$\begin{aligned} \frac{\partial l(\boldsymbol{\theta}')}{\partial \boldsymbol{\phi}'} &= -\frac{1}{N_0} \int_0^T 2 (r_0(t) - \mathbf{h}^T(t) \boldsymbol{\phi}') (-\mathbf{h}^T(t)) dt \\ &\quad - \frac{1}{N_0} \int_0^T \sum_{i=1}^{M-1} 2 (r_i(t) - A'_i \mathbf{h}^T(t - \tau'_i) \boldsymbol{\phi}') (-A'_i \mathbf{h}^T(t - \tau'_i)) dt \end{aligned}$$

In order to find the maximum, we equate the above partial derivative to zero, which gives,

$$\begin{aligned} \int_0^T r_0(t) \mathbf{h}^T(t) dt - \boldsymbol{\phi}'^T \left(\int_0^T \mathbf{h}(t) \mathbf{h}^T(t) dt \right) + \sum_{i=1}^{M-1} A'_i \left(\int_0^T r_i(t) \mathbf{h}^T(t - \tau'_i) dt \right) - \\ A_i'^2 \boldsymbol{\phi}'^T \left(\int_0^T \mathbf{h}(t - \tau'_i) \mathbf{h}^T(t - \tau'_i) dt \right) = \mathbf{0} \end{aligned}$$

Using the properties of the vector $\mathbf{h}(t)$ as shown in Appendix B.3, we have

$$\begin{aligned} \int_0^T r_0(t) \mathbf{h}^T(t) dt - (T/2) \boldsymbol{\phi}'^T + \\ \sum_{i=1}^{M-1} A'_i \left(\int_0^T r_i(t) \mathbf{h}^T(t - \tau'_i) dt \right) - (T/2) \sum_{i=1}^{M-1} A_i'^2 \boldsymbol{\phi}'^T = \mathbf{0} \end{aligned}$$

If we replace the integrals with

$$\mathbf{y}'_0 = \int_0^T r_0(t) \mathbf{h}(t) dt \quad \text{and} \quad \mathbf{y}'_i = \int_0^T r_i(t) \mathbf{h}(t - \tau'_i) dt \quad \text{for } i = 1, 2, \dots, M-1,$$

then we have

$$\mathbf{y}'_0{}^T - (T/2) \boldsymbol{\phi}'^T + \sum_{i=1}^{M-1} A'_i \mathbf{y}'_i{}^T - (T/2) \sum_{i=1}^{M-1} A_i'^2 \boldsymbol{\phi}'^T = \mathbf{0}$$

So, the MLE of ϕ' is

$$\hat{\phi}' = \frac{(2/T) \left(\mathbf{y}'_0 + \sum_{i=1}^{M-1} A'_i \mathbf{y}'_i \right)}{(1 + \mathbf{A}'^T \mathbf{A}')}.$$

Putting this back in (B.23), we have

$$l(\boldsymbol{\theta}') = -\frac{1}{N_0} \sum_{i=0}^{M-1} \int_0^T x_i^2(t) dt + \frac{1}{N_0} \frac{(2/T) \left(\mathbf{y}'_0{}^T + \sum_{i=1}^{M-1} A'_i \mathbf{y}'_i{}^T \right) \left(\mathbf{y}'_0 + \sum_{i=1}^{M-1} A'_i \mathbf{y}'_i \right)}{(1 + \mathbf{A}'^T \mathbf{A}')}.$$

Maximizing $l(\boldsymbol{\theta}')$ w.r.t \mathbf{A}' and $\boldsymbol{\tau}'$ is equivalent to maximizing the second term. So, let

$$f(\mathbf{A}', \boldsymbol{\tau}') = \frac{\left(\mathbf{y}'_0{}^T + \sum_{i=1}^{M-1} A'_i \mathbf{y}'_i{}^T \right) \left(\mathbf{y}'_0 + \sum_{i=1}^{M-1} A'_i \mathbf{y}'_i \right)}{(1 + \mathbf{A}'^T \mathbf{A}')}.$$

If we let $\mathbf{Y}' = [\mathbf{y}'_0 \mathbf{y}'_1 \cdots \mathbf{y}'_{M-1}]$ be the $2N-1 \times M$ matrix then the maximum value of $f(\mathbf{A}', \boldsymbol{\tau}')$ w.r.t \mathbf{A}' is $f_{\max}(\boldsymbol{\tau}')$ is equal to the maximum eigenvalue of $\mathbf{Y}'\mathbf{Y}'^T$. Let $\mathbf{B}' = \mathbf{Y}'\mathbf{Y}'^T$. The matrix \mathbf{B}' is a function of $\boldsymbol{\tau}' = \mathbf{g}'(\boldsymbol{\eta}')$ which is a function of the emitter location (x_T, y_T) . So, the MLE of the emitter location is found by maximizing the maximum eigenvalue of $\mathbf{B}'(x_T, y_T)$. i.e,

$$(\hat{x}_T, \hat{y}_T) = \arg \max_{(x_T, y_T)} \lambda_{\max}(\mathbf{B}'(x_T, y_T)) \quad (\text{B.24})$$

B.2.2 Signal known with unknown transmission time

The log-likelihood function is given by

$$l(\boldsymbol{\zeta}) = -\frac{1}{N_0} \int_0^T \sum_{i=0}^{M-1} (r_i(t) - A_i \mathbf{h}^T(t - \tau_i) \boldsymbol{\phi})^2 dt \quad (\text{B.20})$$

Partial differentiation w.r.t A_k gives,

$$\frac{\partial l(\boldsymbol{\theta})}{\partial A_k} = -\frac{1}{N_0} \int_0^T 2 (r_k(t) - A_k \mathbf{h}^T(t - \tau_k) \boldsymbol{\phi}) (-\mathbf{h}^T(t - \tau_k) \boldsymbol{\phi}) dt = 0$$

for each of $k = 0, 1, \dots, M-1$. Equating this to zero to find the maximum value gives,

$$\boldsymbol{\phi}^T \left(\int_0^T r_k(t) \mathbf{h}(t - \tau_k) dt \right) - A_k \boldsymbol{\phi}^T \left(\int_0^T \mathbf{h}(t - \tau_k) \mathbf{h}^T(t - \tau_k) dt \right) \boldsymbol{\phi} = 0$$

If we replace the integral with $\mathbf{y}_k = \int_0^T r_k(t) \mathbf{h}(t - \tau_k) dt$ and use the properties of the vector $\mathbf{h}(t)$ as shown in Appendix B.3, we have

$$\boldsymbol{\phi}^T \mathbf{y}_k - (T/2) A_k \boldsymbol{\phi}^T \boldsymbol{\phi} = 0$$

So, the MLE of A_k is

$$\hat{A}_k = \frac{\boldsymbol{\phi}^T \mathbf{y}_k}{(T/2)\boldsymbol{\phi}^T \boldsymbol{\phi}}, \quad k = 0, 1, \dots, M-1$$

Putting this back in equation (B.20), we have

$$\begin{aligned} l(\boldsymbol{\zeta}) &= -\frac{1}{N_0} \sum_{i=0}^{M-1} \int_0^T r_i^2(t) dt - 2 \left(\frac{\boldsymbol{\phi}^T \mathbf{y}_i}{(T/2)\boldsymbol{\phi}^T \boldsymbol{\phi}} \right) \left(\int_0^T r_i(t) \mathbf{h}^T(t - \tau_i) dt \right) \boldsymbol{\phi} + \\ &\quad \left(\frac{\boldsymbol{\phi}^T \mathbf{y}_i}{(T/2)\boldsymbol{\phi}^T \boldsymbol{\phi}} \right)^2 \boldsymbol{\phi}^T \left(\int_0^T \mathbf{h}(t - \tau_i) \mathbf{h}^T(t - \tau_i) dt \right) \boldsymbol{\phi} \\ &= -\frac{1}{N_0} \sum_{i=0}^{M-1} \int_0^T r_i^2(t) dt + \frac{1}{N_0} \sum_{i=0}^{M-1} \left(\frac{\boldsymbol{\phi}^T \mathbf{y}_i \mathbf{y}_i^T \boldsymbol{\phi}}{(T/2)\boldsymbol{\phi}^T \boldsymbol{\phi}} \right) \end{aligned}$$

Maximizing $l(\boldsymbol{\zeta})$ w.r.t $\boldsymbol{\tau}$ is equivalent to maximizing the second term. So the MLE for $\boldsymbol{\eta}$ is given by

$$\hat{\boldsymbol{\eta}} = \arg \max_{\boldsymbol{\eta}} \sum_{i=0}^{M-1} \boldsymbol{\phi}^T \mathbf{y}_i \mathbf{y}_i^T \boldsymbol{\phi} = \arg \max_{\boldsymbol{\eta}} \boldsymbol{\phi}^T \mathbf{B} \boldsymbol{\phi} \quad (\text{B.25})$$

where $\mathbf{B} = \mathbf{Y}\mathbf{Y}^T$ and $\mathbf{Y} = [\mathbf{y}_0 \ \mathbf{y}_1 \ \dots \ \mathbf{y}_{M-1}]$ with $\mathbf{y}_i = \int_0^T r_i(t) \mathbf{h}(t - \tau_i) dt$, $i = 0, 1, \dots, M-1$. \mathbf{B} is a function of (x_T, y_T, t_0) .

B.3 Properties of $\mathbf{h}(t)$

The time dependent vector $\mathbf{h}(t)$ that was used for modeling the problem in equation (21) has some interesting properties which simplifies the derivation of the CRLB and the MLE. These properties are derived here. We have

$$\begin{aligned} \mathbf{h}(t - \tau_i) &= \left[\frac{1}{\sqrt{2}} \cos 2\pi F_0(t - \tau_i) \ \dots \ \cos 2\pi(N-1)F_0(t - \tau_i) \ \sin 2\pi F_0(t - \tau_i) \right. \\ &\quad \left. \dots \ \sin 2\pi(N-1)F_0(t - \tau_i) \right]^T \end{aligned}$$

Differentiating both sides w.r.t τ_i , we get

$$\frac{\partial \mathbf{h}(t - \tau_i)}{\partial \tau_i} = 2\pi F_0 \begin{bmatrix} 0 & \sin 2\pi F_0(t - \tau_i) & 2\sin 2\pi F_0(t - \tau_i) & \dots & (N-1)\sin 2\pi(N-1)F_0(t - \tau_i) \\ -\cos 2\pi F_0(t - \tau_i) & -2\cos 2\pi F_0(t - \tau_i) & \dots & -(N-1)\cos 2\pi(N-1)F_0(t - \tau_i) \end{bmatrix}^T$$

Let

$$\mathbf{L} = \begin{bmatrix} \mathbf{0}_{(N,N)} & \begin{bmatrix} \mathbf{0}_{(1,N-1)} \\ \text{diag}(1, 2, \dots, N-1) \end{bmatrix} \\ -\begin{bmatrix} \mathbf{0}_{(N-1,1)} & \text{diag}(1, 2, \dots, N-1) \end{bmatrix} & \mathbf{0}_{(N-1,N-1)} \end{bmatrix}$$

So, we have the partial derivative of $\mathbf{h}(t - \tau_i)$ w.r.t to τ_i as

$$\frac{\partial \mathbf{h}(t - \tau_i)}{\partial \tau_i} = (2\pi F_0) \mathbf{L} \mathbf{h}(t - \tau_i)$$

Next we will compute the integral $\int_0^T \mathbf{h}(t)\mathbf{h}^T(t) dt$. We have,

$$\mathbf{h}(t)\mathbf{h}^T(t) = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \cos 2\pi F_0 t \\ \vdots \\ \cos 2\pi(N-1)F_0 t \\ \sin 2\pi F_0 t \\ \vdots \\ \sin 2\pi(N-1)F_0 t \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \cos 2\pi F_0 t \\ \vdots \\ \cos 2\pi(N-1)F_0 t \\ \sin 2\pi F_0 t \\ \vdots \\ \sin 2\pi(N-1)F_0 t \end{bmatrix}^T$$

Let us compute each of the integrals in this $2N-1 \times 2N-1$ product matrix separately. Integral of the first element is

$$\int_0^T \frac{1}{2} dt = [t]_0^T = \frac{T}{2}$$

Integrals of the elements on the diagonal are given by

$$\begin{aligned} \int_0^T \cos^2 2\pi k F_0 t dt &= \frac{1}{2} \int_0^T (1 + \cos 4\pi k F_0 t) dt \\ &= \frac{1}{2} \left[t + \frac{\sin 4\pi k F_0 t}{4\pi k F_0} \right]_0^T \\ &= \frac{T}{2} \end{aligned}$$

and

$$\begin{aligned} \int_0^T \sin^2 2\pi k F_0 t dt &= \frac{1}{2} \int_0^T (1 - \cos 4\pi k F_0 t) dt \\ &= \frac{1}{2} \left[t - \frac{\sin 4\pi k F_0 t}{4\pi k F_0} \right]_0^T \\ &= \frac{T}{2} \end{aligned}$$

for $k = 1, 2, \dots, N-1$. Integrals of the rest of the elements are given by

$$\begin{aligned} \int_0^T \cos 2\pi k F_0 t \sin 2\pi n F_0 t dt &= \frac{1}{2} \int_0^T \sin 2\pi(n+k)F_0 t dt + \sin 2\pi(n-k)F_0 t dt \\ &= \frac{1}{2} \left[-\frac{\cos 2\pi(n+k)F_0 t}{4\pi(n+k)F_0} - \frac{\cos 2\pi(n-k)F_0 t}{4\pi(n-k)F_0} \right]_0^T \\ &= 0 \end{aligned}$$

and

$$\begin{aligned} \int_0^T \cos 2\pi k F_0 t \cos 2\pi n F_0 t dt &= \frac{1}{2} \int_0^T \cos 2\pi(n+k)F_0 t dt + \cos 2\pi(n-k)F_0 t dt \\ &= \frac{1}{2} \left[\frac{\sin 2\pi(n+k)F_0 t}{4\pi(n+k)F_0} + \frac{\sin 2\pi(n-k)F_0 t}{4\pi(n-k)F_0} \right]_0^T \\ &= 0 \end{aligned}$$

and

$$\begin{aligned}
\int_0^T \sin 2\pi k F_0 t \sin 2\pi n F_0 t \, dt &= \frac{1}{2} \int_0^T \cos 2\pi(k-n)F_0 t \, dt - \cos 2\pi(n+k)F_0 t \, dt \\
&= \frac{1}{2} \left[\frac{\sin 2\pi(k-n)F_0 t}{4\pi(k-n)F_0} + \frac{\sin 2\pi(n+k)F_0 t}{4\pi(n+k)F_0} \right]_0^T \\
&= 0
\end{aligned}$$

for $k, n = 1, 2, \dots, N-1$. Therefore, we have the integral of $\mathbf{h}(t)\mathbf{h}^T(t)$ as a scaled identity matrix given by

$$\int_0^T \mathbf{h}(t)\mathbf{h}^T(t) \, dt = \frac{T}{2} \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = (T/2)\mathbf{I}_{(2N-1)}$$

Since $\mathbf{h}(t)$ is periodic with period T , for any τ ,

$$\int_0^T \mathbf{h}(t-\tau)\mathbf{h}^T(t-\tau) \, dt = \int_0^T \mathbf{h}(t)\mathbf{h}^T(t) \, dt = (T/2)\mathbf{I}_{(2N-1)} \quad (\text{B.26})$$

Now, we compute the integral of the cross-product of the partial derivatives of $\mathbf{h}(t-\tau_i)$ w.r.t to τ_i

$$\begin{aligned}
\int_0^T \frac{\partial \mathbf{h}(t-\tau_i)}{\partial \tau_i} \frac{\partial \mathbf{h}^T(t-\tau_i)}{\partial \tau_i} \, dt &= \int_0^T (2\pi F_0 \mathbf{L} \mathbf{h}(t-\tau_i)) (2\pi F_0 \mathbf{L} \mathbf{h}(t-\tau_i))^T \, dt \\
&= (2\pi F_0)^2 \mathbf{L} \left[\int_0^T \mathbf{h}(t-\tau_i) \mathbf{h}^T(t-\tau_i) \, dt \right] \mathbf{L}^T \\
&= (T/2)(2\pi F_0)^2 \mathbf{L} \mathbf{L}^T
\end{aligned} \quad (\text{B.27})$$

and the integral of the cross-product of $\mathbf{h}(t-\tau_i)$ with its partial derivative w.r.t to τ_i is

$$\begin{aligned}
\int_0^T \frac{\partial \mathbf{h}(t-\tau_i)}{\partial \tau_i} \mathbf{h}^T(t-\tau_i) \, dt &= \int_0^T 2\pi F_0 \mathbf{L} \mathbf{h}(t-\tau_i) \mathbf{h}^T(t-\tau_i) \, dt \\
&= (2\pi F_0) \mathbf{L} \left[\int_0^T \mathbf{h}(t-\tau_i) \mathbf{h}^T(t-\tau_i) \, dt \right] \\
&= (T/2)(2\pi F_0) \mathbf{L}
\end{aligned} \quad (\text{B.28})$$

B.4 Transformation of the Parameters

In Section 3.3 we discussed the relationship between the unknown attenuation factors and the unknown signal, and between the unknown TOAs and the unknown

signal. Here we will show that the FIM given in equation (24) is rank two deficient. Then we will show that the transformation given in equation (25) satisfies the conditions given in [40]. From (24), we have

$$\mathcal{I}_\theta = \frac{(T/2)}{(N_0/2)} \begin{bmatrix} (2\pi F_0)^2 \phi^T \mathbf{L} \mathbf{L}^T \phi (\text{diag}(\mathbf{A}))^2 & (2\pi F_0)(\phi^T \mathbf{L} \phi)(\text{diag}(\mathbf{A})) & (2\pi F_0)(\mathbf{A} \odot \mathbf{A}) \phi^T \mathbf{L} \\ (2\pi F_0)(\phi^T \mathbf{L}^T \phi)(\text{diag}(\mathbf{A})) & (\phi^T \phi) \mathbf{I}_M & \mathbf{A} \phi^T \\ (2\pi F_0) \mathbf{L}^T \phi (\mathbf{A} \odot \mathbf{A})^T & \phi \mathbf{A}^T & (\mathbf{A}^T \mathbf{A}) \mathbf{I}_{(2N-1)} \end{bmatrix}$$

If $\boldsymbol{\nu}_1 = [\mathbf{1}_M^T \quad \mathbf{A}^T \quad -(\phi + (2\pi F_0) \mathbf{L}^T \phi)]^T$ and $\boldsymbol{\nu}_2 = [\mathbf{1}_M^T \quad -\mathbf{A}^T \quad (\phi - (2\pi F_0) \mathbf{L}^T \phi)]^T$ then it can be verified that $\mathcal{I}_\theta \boldsymbol{\nu}_1 = \mathbf{0}$ and $\mathcal{I}_\theta \boldsymbol{\nu}_2 = \mathbf{0}$. Therefore $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ are in the null space of \mathcal{I}_θ . Also, $\mathcal{I}_\theta + (1/2) \boldsymbol{\nu}_1 \boldsymbol{\nu}_1^T + (1/2) \boldsymbol{\nu}_2 \boldsymbol{\nu}_2^T$ is non-singular. This means that $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ are the basis vectors for the null space of \mathcal{I}_θ and so the matrix \mathcal{I}_θ is rank two deficient. The Jacobian of the transformation is given by,

$$\mathbf{H} = \left(\frac{\partial \boldsymbol{\theta}'}{\partial \boldsymbol{\theta}} \right) = \begin{bmatrix} \frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\tau}} & \frac{\partial \boldsymbol{\tau}'}{\partial \mathbf{A}} & \frac{\partial \boldsymbol{\tau}'}{\partial \phi} \\ \frac{\partial \mathbf{A}'}{\partial \boldsymbol{\tau}} & \frac{\partial \mathbf{A}'}{\partial \mathbf{A}} & \frac{\partial \mathbf{A}'}{\partial \phi} \\ \frac{\partial \phi'}{\partial \boldsymbol{\tau}} & \frac{\partial \phi'}{\partial \mathbf{A}} & \frac{\partial \phi'}{\partial \phi} \end{bmatrix} \quad (\text{B.29})$$

Now, we will compute each of the derivatives in the Jacobian matrix. The elements of the first sub-column are given by,

$$\begin{aligned} \frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\tau}} &= \frac{\partial}{\partial \boldsymbol{\tau}} ([-\mathbf{1}_{M-1} \quad \mathbf{I}_{M-1}] \boldsymbol{\tau}) \\ &= [-\mathbf{1}_{M-1} \quad \mathbf{I}_{M-1}] \\ \frac{\partial \boldsymbol{\tau}'}{\partial \mathbf{A}} &= \mathbf{0}_{(M-1, M)} \\ \frac{\partial \boldsymbol{\tau}'}{\partial \phi} &= \mathbf{0}_{(M-1, 2N-1)} \end{aligned}$$

The elements of the second sub-column are given by,

$$\begin{aligned} \frac{\partial \mathbf{A}'}{\partial \boldsymbol{\tau}} &= \mathbf{0}_{(M-1, M)} \\ \frac{\partial \mathbf{A}'}{\partial \mathbf{A}} &= \frac{\partial}{\partial \mathbf{A}} \left(([\mathbf{0}_{(M-1, 1)} \quad \mathbf{I}_{M-1}] \mathbf{A}) (\mathbf{e}_1^T \mathbf{A})^{-1} \right) \\ &= (1/A_0) [-\mathbf{A}' \quad \mathbf{I}_{M-1}] \\ \frac{\partial \mathbf{A}'}{\partial \phi} &= \mathbf{0}_{(M-1, 2N-1)} \end{aligned}$$

where \mathbf{e}_1 is the first column of an $M \times M$ identity matrix. The elements of the third sub-column are given by,

$$\begin{aligned}
\frac{\partial \phi'}{\partial \tau} &= \frac{\partial}{\partial \tau} A_0 \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}(\mathbf{h}(-\tau_0))\phi \\
&= \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}\left(\frac{\partial \mathbf{h}(-\tau_0)}{\partial \tau_0}\right)\phi \quad \mathbf{0}_{(2N-1,1)} \quad \cdots \quad \mathbf{0}_{(2N-1,1)} \\
&= \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} (2\pi F_0) \mathbf{L} \text{diag}(\mathbf{h}(-\tau_0))\phi \quad \mathbf{0}_{(2N-1,1)} \quad \cdots \quad \mathbf{0}_{(2N-1,1)} \\
\frac{\partial \phi'}{\partial \mathbf{A}} &= \frac{\partial}{\partial \mathbf{A}} \left(\begin{bmatrix} \mathbf{e}_1^T \mathbf{A} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}(\mathbf{h}(-\tau_0))\phi \right) \\
&= \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}(\mathbf{h}(-\tau_0))\phi \quad \mathbf{0}_{(2N-1,1)} \quad \cdots \quad \mathbf{0}_{(2N-1,1)} \\
\frac{\partial \phi'}{\partial \phi} &= \frac{\partial}{\partial \phi} \left(\begin{bmatrix} \mathbf{e}_1^T \mathbf{A} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}(\mathbf{h}(-\tau_0))\phi \right) \\
&= \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{N-1} & \mathbf{I}_{N-1} \\ \mathbf{I}_{N-1} & -\mathbf{I}_{N-1} \end{bmatrix} \end{bmatrix} \text{diag}(\mathbf{h}(-\tau_0)) \quad \mathbf{0}_{(2N-1,1)} \quad \cdots \quad \mathbf{0}_{(2N-1,1)}
\end{aligned}$$

For the transformed parameters to have finite variance, the row vectors of \mathbf{H} must be equal to the linear combinations of those eigenvectors of \mathcal{I}_θ that have nonzero eigenvalues [40]. In order to show that the row vectors of \mathbf{H} are linear combinations of those eigenvectors of \mathcal{I}_θ that have nonzero eigenvalues, it is enough to show that the row vectors of \mathbf{H} are orthogonal to the null space of \mathcal{I}_θ . That is, it is enough to show that $\mathbf{H}\nu_1 = \mathbf{0}$ and $\mathbf{H}\nu_2 = \mathbf{0}$. Now,

$$\mathbf{H}\nu_1 = \begin{bmatrix} \frac{\partial \tau'}{\partial \tau} \mathbf{1}_M + \frac{\partial \tau'}{\partial \mathbf{A}} \mathbf{A} - \frac{\partial \tau'}{\partial \phi} (\phi + (2\pi F_0) \mathbf{L}^T \phi) \\ \frac{\partial \mathbf{A}'}{\partial \tau} \mathbf{1}_M + \frac{\partial \mathbf{A}'}{\partial \mathbf{A}} \mathbf{A} - \frac{\partial \mathbf{A}'}{\partial \phi} (\phi + (2\pi F_0) \mathbf{L}^T \phi) \\ \frac{\partial \phi'}{\partial \tau} \mathbf{1}_M + \frac{\partial \phi'}{\partial \mathbf{A}} \mathbf{A} - \frac{\partial \phi'}{\partial \phi} (\phi + (2\pi F_0) \mathbf{L}^T \phi) \end{bmatrix}.$$

Substituting the partial derivatives that we computed previously and further simplifying gives,

$$\mathbf{H}\nu_1 = \begin{bmatrix} [-\mathbf{1}_{M-1} \quad \mathbf{I}_{(M-1)}] \mathbf{1}_M + \mathbf{0}_{(M-1,1)} - \mathbf{0}_{(M-1,1)} \\ \mathbf{0}_{(M-1,1)} + (1/A_0) [-\mathbf{A}' \quad \mathbf{I}_{(M-1)}] \mathbf{A} - \mathbf{0}_{(M-1,1)} \\ A_0 \mathbf{P} (2\pi F_0) \mathbf{L} \text{diag}(\mathbf{h}(-\tau_0))\phi + A_0 \mathbf{P} \text{diag}(\mathbf{h}(-\tau_0))\phi - \\ (A_0 \mathbf{P} \text{diag}(\mathbf{h}(-\tau_0))\phi + A_0 \mathbf{P} \text{diag}(\mathbf{h}(-\tau_0))(2\pi F_0) \mathbf{L}^T \phi) \end{bmatrix}$$

where

$$\mathbf{P} = \begin{bmatrix} 1 & \mathbf{0}_{(1,2N-2)} \\ \mathbf{0}_{(2N-2,1)} & \begin{bmatrix} \mathbf{I}_{(N-1)} & \mathbf{I}_{(N-1)} \\ \mathbf{I}_{(N-1)} & -\mathbf{I}_{(N-1)} \end{bmatrix} \end{bmatrix}.$$

Using the fact that $\mathbf{L} \text{diag}(\mathbf{h}(-\tau_0)) = \text{diag}(\mathbf{h}(-\tau_0)) \mathbf{L}^T$, we have $\mathbf{H}\nu_1 = \mathbf{0}$. Similarly it can be shown that $\mathbf{H}\nu_2 = \mathbf{0}$.

C Optimal Sensor Configuration

Here we will analyze the effect of the sensor-emitter geometry on the FIM for the emitter location. We will derive the D-optimal configurations for the unconstrained and the constrained cases that are described in Section 4.1. From (30) and (B.18) we have the FIM for the emitter location vector $\boldsymbol{\eta}' = [x_T \ y_T]$ as

$$\mathcal{I}_{\eta'} = \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T (2\pi F_0)^2 \boldsymbol{\phi}'^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi}' (\text{diag}(\mathbf{A}'))^2 \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) \quad (\text{C.1})$$

If we assume that the signal level is the same at all sensors, i.e, $A_i = A$ for all i , then we have $\text{diag}(\mathbf{A}') = A \mathbf{I}_{(M-1)}$, where $\mathbf{I}_{(M-1)}$ is the $(M-1) \times (M-1)$ identity matrix, and we have the above FIM as

$$\mathcal{I}_{\eta'} = (2\pi F_0)^2 \boldsymbol{\phi}'^T \mathbf{L} \mathbf{L}^T \boldsymbol{\phi}' A^2 \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) \quad (\text{C.2})$$

If we let ψ_i be the angle of sensor i from the positive x-axis measured at the emitter as shown in Figure (D.1), then we can write (B.19) as

$$\left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) = (1/c) \begin{bmatrix} (\cos \psi_1 - \cos \psi_0) & (\sin \psi_1 - \sin \psi_0) \\ (\cos \psi_2 - \cos \psi_0) & (\sin \psi_2 - \sin \psi_0) \\ \vdots & \vdots \\ (\cos \psi_{M-1} - \cos \psi_0) & (\sin \psi_{M-1} - \sin \psi_0) \end{bmatrix} \quad (\text{C.3})$$

Now let

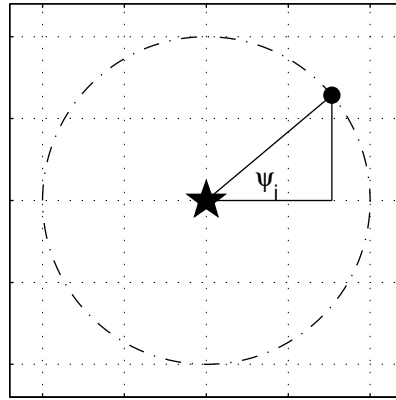


Figure C.1. Definition of the angle ψ_i .

$$\mathbf{J} = \begin{bmatrix} \cos \psi_0 & \cos \psi_1 & \cdots & \cos \psi_{M-1} \\ \sin \psi_0 & \sin \psi_1 & \cdots & \sin \psi_{M-1} \\ 1 & 1 & \cdots & 1 \end{bmatrix}$$

Using elementary operations, we can transform \mathbf{J} to

$$\begin{bmatrix} \cos \psi_0 & \cos \psi_1 - \cos \psi_0 & \cdots & \cos \psi_{M-1} - \cos \psi_0 \\ \sin \psi_0 & \sin \psi_1 - \sin \psi_0 & \cdots & \sin \psi_{M-1} - \sin \psi_0 \\ 1 & 0 & \cdots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{u} & c \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \\ 1 & \mathbf{0}^T \end{bmatrix}$$

where $\mathbf{u} = [\cos \psi_0 \quad \sin \psi_0 \quad 1]^T$ and $\mathbf{0}$ is the $(M-1) \times 1$ zero vector. Now, since the elementary operations do not change the determinant, we have

$$\begin{aligned} \det(\mathbf{J}\mathbf{J}^T) &= \det \left\{ \begin{bmatrix} \mathbf{u} & c \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \\ 1 & \mathbf{0}^T \end{bmatrix} \begin{bmatrix} \mathbf{u}^T & 1 \\ c \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) & \mathbf{0} \end{bmatrix} \right\} \\ &= \det \left\{ \begin{bmatrix} \mathbf{u}\mathbf{u}^T + c^2 \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) & \mathbf{u} \\ \mathbf{u}^T & 1 \end{bmatrix} \right\} \\ &= \det \left(\mathbf{u}\mathbf{u}^T + c^2 \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) - \mathbf{u}\mathbf{u}^T \right) \end{aligned}$$

So, $\det \left[\left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right)^T \left(\frac{\partial \boldsymbol{\tau}'}{\partial \boldsymbol{\eta}'^T} \right) \right] = (1/c^2)^3 \det(\mathbf{J}\mathbf{J}^T)$ so that we have,

$$\det(\mathcal{I}_{\eta'}) = \frac{(2\pi F_0)^2 \boldsymbol{\phi}'^T \mathbf{L}\mathbf{L}^T \boldsymbol{\phi}'}{c^6} \det(\mathbf{J}\mathbf{J}^T)$$

Maximizing the determinant of the FIM is equivalent to maximizing $\det(\mathbf{J}\mathbf{J}^T)$.

C.1 Unconstrained Geometry

This problem can be stated as follows:

We have $g : \Re^M \rightarrow \Re$ such that

$$g(\boldsymbol{\psi}) = \det(\mathbf{J}\mathbf{J}^T)$$

where $\boldsymbol{\psi} = [\psi_0 \quad \psi_1 \quad \cdots \quad \psi_{M-1}]^T$, $\Psi = [-\pi, \pi]^M$ is a convex subset of \Re^M . Find

$$\arg \max_{\boldsymbol{\psi} \in \Psi} g(\boldsymbol{\psi})$$

We have,

$$\mathbf{J}\mathbf{J}^T = \begin{bmatrix} \sum_{i=0}^{M-1} \cos^2 \psi_i & \sum_{i=0}^{M-1} \cos \psi_i \sin \psi_i & \sum_{i=0}^{M-1} \cos \psi_i \\ \sum_{i=0}^{M-1} \cos \psi_i \sin \psi_i & \sum_{i=0}^{M-1} \sin^2 \psi_i & \sum_{i=0}^{M-1} \sin \psi_i \\ \sum_{i=0}^{M-1} \cos \psi_i & \sum_{i=0}^{M-1} \sin \psi_i & M \end{bmatrix}$$

By Hadamard's inequality,

$$g(\boldsymbol{\psi}) = \det(\mathbf{J}\mathbf{J}^T) \leq \left(\sum_{i=0}^{M-1} \cos^2 \psi_i \right) \left(\sum_{i=0}^{M-1} \sin^2 \psi_i \right) (M)$$

The equality holds if and only if $\mathbf{J}\mathbf{J}^T$ is diagonal. That is iff

$$\begin{aligned} \sum_{i=0}^{M-1} \cos \psi_i \sin \psi_i &= 0 \\ \sum_{i=0}^{M-1} \cos \psi_i &= 0 \\ \sum_{i=0}^{M-1} \sin \psi_i &= 0 \end{aligned} \tag{C.4}$$

Also,

$$\begin{aligned} \left(\sum_{i=0}^{M-1} \cos^2 \psi_i \right) \left(\sum_{i=0}^{M-1} \sin^2 \psi_i \right) &= \left(\sum_{i=0}^{M-1} \left(\frac{1}{2} + \frac{1}{2} \cos 2\psi_i \right) \right) \left(\sum_{i=0}^{M-1} \left(\frac{1}{2} - \frac{1}{2} \cos 2\psi_i \right) \right) \\ &= \left(\frac{M}{2} + \frac{1}{2} \sum_{i=0}^{M-1} \cos 2\psi_i \right) \left(\frac{M}{2} - \frac{1}{2} \sum_{i=0}^{M-1} \cos 2\psi_i \right) \\ &= \frac{M^2}{4} - \frac{1}{4} \left(\sum_{i=0}^{M-1} \cos 2\psi_i \right)^2 \leq \frac{M^2}{4} \end{aligned}$$

This inequality holds iff

$$\sum_{i=0}^{M-1} \cos 2\psi_i = 0 \tag{C.5}$$

When the sensors are placed around the emitter at equi-angular distances (as shown in Figure 17(a)), then we have $\psi_i = 2\pi \left(\frac{i}{M} \right)$, $i = 0, \dots, M-1$. This configuration satisfies the conditions in (C.4) and (C.5). Therefore, this is an optimal configuration. Also, when M is a multiple of three, if the sensors are distributed equally at the three vertices of an equilateral triangle inscribed in the circle around the emitter, as shown in Figure 17(b), then this configuration also satisfy these conditions and hence, is also an optimal configuration.

C.2 Constrained Geometry

This problem can be stated as follows:

We have $g : \Re^M \rightarrow \Re$ such that

$$g(\boldsymbol{\psi}) = \det(\mathbf{J}\mathbf{J}^T)$$

where $\boldsymbol{\psi} = [\psi_0 \ \psi_1 \ \dots \ \psi_{M-1}]^T$, $\Psi = [-\theta, \theta]^M$ is a convex subset of \Re^M . Find

$$\arg \max_{\boldsymbol{\psi} \in \Psi} g(\boldsymbol{\psi})$$

Notice that here the domain is constrained. For this constrained geometry problem we will derive the D-optimal configurations for three and four sensors.

Three sensor case

Here we use a simple transformation of variables. Let $\cos \psi_i = x_i$, $i = 0, 1, 2$, so that $\sin \psi_i = \sqrt{1 - x_i^2}$. With this transformation of variables, the problem can be restated as:

Suppose that $\mathbf{x} = [x_0 \ x_1 \ x_2]$ and $X = [-\zeta, \zeta] \times [x_0, \zeta] \times [x_1, \zeta] \subset \mathbb{R}^3$. Let

$$\mathbf{J} = \begin{bmatrix} x_0 & x_1 & x_2 \\ f(x_0) & f(x_1) & f(x_2) \\ 1 & 1 & 1 \end{bmatrix}$$

where $f(x) = \sqrt{1 - x^2}$ and let $g : \mathbb{R}^3 \rightarrow \mathbb{R}$ such that $g(\mathbf{x}) = \det(\mathbf{J}\mathbf{J}^T) = (\det(\mathbf{J}))^2$. Find

$$\arg \max_{\mathbf{x} \in X} g(\mathbf{x})$$

The setup is as shown in Figure C.2.

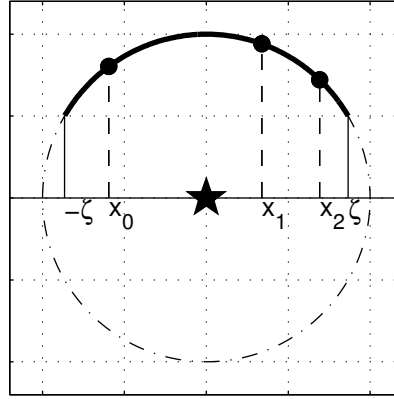


Figure C.2. Three sensor setup.

Now, for arbitrary $x_1 \geq -\zeta$, we have $x_0 \in [-\zeta, x_1]$. We will show that for all arbitrary fixed $x_1, x_2 \in X$, $g(x_0)$ is a convex function of x_0 for $x_0 \in [-\zeta, x_1]$. We have

$$\frac{\partial g(x_0)}{\partial x_0} = \frac{\partial}{\partial x_0} (\det(\mathbf{J}))^2 = 2 \det(\mathbf{J}) \frac{\partial}{\partial x_0} \det(\mathbf{J})$$

Now,

$$\det(\mathbf{J}) = x_0(f(x_1) - f(x_2)) - f(x_0)(x_1 - x_2) + (x_1 f(x_2) - x_2 f(x_1))$$

\Rightarrow

$$\frac{\partial}{\partial x_0} \det(\mathbf{J}) = (f(x_1) - f(x_2)) - f'(x_0)(x_1 - x_2)$$

\Rightarrow

$$\frac{\partial g(x_0)}{\partial x_0} = 2 \det(\mathbf{J}) ((f(x_1) - f(x_2)) - f'(x_0)(x_1 - x_2))$$

\Rightarrow

$$\frac{\partial^2 g(x_0)}{\partial x_0^2} = 2 \left(\frac{\partial \det(\mathbf{J})}{\partial x_0} \right)^2 + 2 \det(\mathbf{J}) \frac{\partial^2 \det(\mathbf{J})}{\partial x_0^2}$$

Now

$$\frac{\partial \det(\mathbf{J})}{\partial x_0} = (f(x_1) - f(x_2)) - f'(x_0)(x_1 - x_2)$$

\Rightarrow

$$\frac{\partial^2 \det(\mathbf{J})}{\partial x_0^2} = -f''(x_0)(x_1 - x_2) < 0$$

since f is concave and $x_1 < x_2$. Next we will show that $\det(\mathbf{J}) < 0$

$$\begin{aligned} \det(\mathbf{J}) &= x_0(f(x_1) - f(x_2)) - f(x_0)(x_1 - x_2) + (x_1 f(x_2) - x_2 f(x_1)) \\ &= f(x_0)(x_2 - x_1) + f(x_2)(x_1 - x_0) - f(x_1)(x_2 - x_0) \end{aligned}$$

Choose $\alpha, \beta \in (0, 1)$ such that $\alpha + \beta = 1$ and $\alpha x_0 + \beta x_2 = x_1$. So, we get $\alpha = (x_2 - x_1)/(x_2 - x_0)$ and $\beta = (x_1 - x_0)/(x_2 - x_0)$. Now, since f is concave,

$$f(x_1) = f(\alpha x_0 + \beta x_2) > \alpha f(x_0) + \beta f(x_2) = \frac{(x_2 - x_1)f(x_0)}{(x_2 - x_0)} + \frac{(x_1 - x_0)f(x_2)}{(x_2 - x_0)}$$

\Rightarrow

$$(x_2 - x_0)f(x_1) > f(x_0)(x_2 - x_1) + f(x_2)(x_1 - x_0)$$

Therefore, $\det(\mathbf{J}) < 0$ and so $\frac{\partial^2 g(x_0)}{\partial x_0^2} > 0$. This means that $g(x_0)$ is a convex function of $x_0 \in [-\zeta, x_1]$. The maximum of $g(x_0)$ occurs at $x_0 = -\zeta$ or $x_0 = x_1$. When $x_0 = x_1$, $g(\mathbf{x}) = 0$. So the x_0 that maximizes $g(\mathbf{x})$ is $x_0 = -\zeta$. Similarly, we can show that the x_2 that maximizes $g(\mathbf{x})$ is $x_2 = \zeta$. Now fix $x_0 = -\zeta$ and $x_2 = \zeta$. So $x_1 \in [-\zeta, \zeta]$.

$$\frac{\partial g(\mathbf{x})}{\partial x_1} = 2 \det(\mathbf{J}) \frac{\partial}{\partial x_1} \det(\mathbf{J})$$

Now,

$$\det(\mathbf{J}) = x_0(f(x_1) - f(x_2)) - f(x_0)(x_1 - x_2) + (x_1 f(x_2) - x_2 f(x_1))$$

\Rightarrow

$$\frac{\partial}{\partial x_1} \det(\mathbf{J}) = (f(x_2) - f(x_0)) + f'(x_1)(x_0 - x_2)$$

\Rightarrow

$$\frac{\partial g(\mathbf{x})}{\partial x_1} = 2 \det(\mathbf{J})((f(x_2) - f(x_0)) + f'(x_1)(x_0 - x_2))$$

Equating this to zero to determine the extrema

$$2 \det(\mathbf{J})((f(x_2) - f(x_0)) + f'(x_1)(x_0 - x_2)) = 0$$

\Rightarrow

$$\det(\mathbf{J}) = 0 \quad \text{or} \quad (f(x_2) - f(x_0)) + f'(x_1)(x_0 - x_2) = 0$$

If $\det(\mathbf{J}) = 0$ then $g(\mathbf{x}) = 0$, so this is not a maximum. When $(f(x_2) - f(x_0)) + f'(x_1)(x_0 - x_2) = 0$, we have

$$f'(x_1) = \frac{(f(x_2) - f(x_0))}{(x_2 - x_0)}$$

\Rightarrow

$$\frac{-x_1}{\sqrt{1-x_1^2}} = \frac{f(\zeta) - f(-\zeta)}{\zeta + \zeta} = 0$$

Therefore $x_1 = 0$. The solution for the problem is

$$\arg \max_{\mathbf{x} \in X} g(\mathbf{x}) = \{[-\zeta \ 0 \ \zeta]^T\}$$

Four sensor case

Here we assume the domain $\Psi = [0, \pi] \times [\psi_0, \pi] \times [\psi_1, \pi] \times [\psi_2, \pi] \subset \Re^4$. Notice that here we are assuming that the constrained arc is a semicircle. The more general case where $\theta \leq \psi_0 \leq \psi_1 \leq \psi_2 \leq \pi - \theta$ where $0 < \theta < \pi/2$ appears to be more difficult. Using the argument from above, we can show that for optimal configuration $\psi_0 = 0$ and $\psi_3 = \pi$. This is shown in the Figure C.3. So, we have

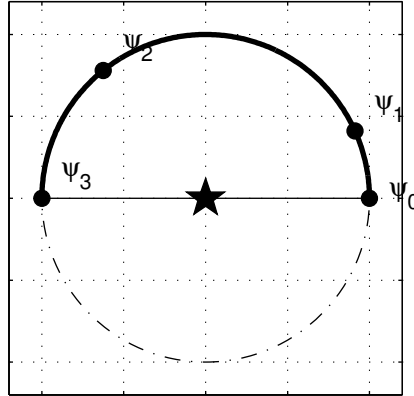


Figure C.3. Four sensor setup.

$$\mathbf{J} = \begin{bmatrix} 1 & \cos \psi_1 & \cos \psi_2 & -1 \\ 0 & \sin \psi_1 & \sin \psi_2 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

We want to prove that $(1/4) \det(\mathbf{J}\mathbf{J}^T) \leq 2$. Note that from Cauchy-Binet theorem

$$\begin{aligned} \det(\mathbf{J}\mathbf{J}^T) &= \det^2 \left(\begin{bmatrix} 1 & \cos \psi_1 & \cos \psi_2 \\ 0 & \sin \psi_1 & \sin \psi_2 \\ 1 & 1 & 1 \end{bmatrix} \right) + \det^2 \left(\begin{bmatrix} 1 & \cos \psi_1 & -1 \\ 0 & \sin \psi_1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \right) + \\ &\quad \det^2 \left(\begin{bmatrix} 1 & \cos \psi_2 & -1 \\ 0 & \sin \psi_2 & 0 \\ 1 & 1 & 1 \end{bmatrix} \right) + \det^2 \left(\begin{bmatrix} \cos \psi_1 & \cos \psi_2 & -1 \\ \sin \psi_1 & \sin \psi_2 & 0 \\ 1 & 1 & 1 \end{bmatrix} \right) \end{aligned}$$

$$\begin{aligned}
&= (\sin \psi_1 + \sin \psi_2 \cos \psi_1 - \sin \psi_1 \cos \psi_2 - \sin \psi_2)^2 + (2 \sin \psi_1)^2 + (2 \sin \psi_2)^2 + \\
&\quad (\sin \psi_2 \cos \psi_1 - \sin \psi_1 + \sin \psi_2 - \sin \psi_1 \cos \psi_2)^2 \\
&= 4 \sin^2 \psi_1 + 4 \sin^2 \psi_2 + 2(\sin \psi_1 - \sin \psi_2)^2 + 2(\sin \psi_2 \cos \psi_1 - \sin \psi_1 \cos \psi_2)^2
\end{aligned}$$

Hence we need to prove that

$$\sin^2 \psi_1 + \sin^2 \psi_2 + (1/2)(\sin \psi_1 - \sin \psi_2)^2 + (1/2)(\sin \psi_2 \cos \psi_1 - \sin \psi_1 \cos \psi_2)^2 \leq 2 \quad (\text{C.6})$$

(C.6) is equivalent to the following inequality, which can be easily verified

$$(1/2)(\sin \psi_2 \cos \psi_2 + \sin \psi_1 \cos \psi_1)^2 + (1/2)(\sin^2 \psi_1 + \sin^2 \psi_2 - 2)^2 \geq (1/2)(\sin \psi_1 - \sin \psi_2)^2 \quad (\text{C.7})$$

Now, if we show that

$$(\sin^2 \psi_1 + \sin^2 \psi_2 - 2)^2 \geq (\sin \psi_1 - \sin \psi_2)^2 \quad (\text{C.8})$$

then (C.7) is proved since $(\sin \psi_2 \cos \psi_2 + \sin \psi_1 \cos \psi_1)^2 \geq 0$. Without loss of generality (due to symmetry of (C.8)), we assume that $\sin \psi_2 \geq \sin \psi_1$. Hence (C.8) is equivalent to

$$2 - \sin^2 \psi_1 - \sin^2 \psi_2 \geq \sin \psi_2 - \sin \psi_1 \quad (\text{C.9})$$

which is equivalent to

$$5/2 - (\sin \psi_1 - 1/2)^2 - (\sin \psi_2 + 1/2)^2 \geq 0 \quad (\text{C.10})$$

Since $(\sin \psi_1 - 1/2)^2 \leq 1/4$ and $(\sin \psi_2 + 1/2)^2 \leq 9/4$, inequality (C.10) is easily verified and this proves inequality (C.6). Equality in (C.10) holds when $\sin \psi_1 = 0$ or 1, $\sin \psi_2 = 1$ under which condition the equality in (C.7) holds. This means that $\sin \psi_1 = 0$, $\sin \psi_2 = 1$ and $\sin \psi_1 = 1$, $\sin \psi_2 = 1$ are optimal solutions. Note that since we have assumed $\sin \psi_2 \geq \sin \psi_1$, we would also have $\sin \psi_1 = 1$, $\sin \psi_2 = 0$ as the optimal solution. Therefore the optimal sensor configurations are given by

$$\arg \max_{\psi \in \Psi} g(\psi) = \{[0 \ 0 \ \pi/2 \ \pi]^T, [0 \ \pi/2 \ \pi/2 \ \pi]^T, [0 \ \pi/2 \ \pi \ \pi]^T\}$$

D Derivation of the Fisher information matrix

Assume a stationary emitter and stationary receivers. Also to simplify the derivation, assume a known signal and known transmission time. We wish to determine the CRLB for the emitter position.

At receiver i , ($i = 1, \dots, M$) we receive

$$r_i(t) = A_i(x_T, y_T)s(t - \tau_i(x_T, y_T)) + w_i(t)$$

where $s(t)$ is known, $A_i(x_T, y_T)$ and $\tau_i(x_T, y_T)$ are known functions of (x_T, y_T) and $w_i(t)$ is white Gaussian noise with the same power spectral density

$$P_W(F) = N_0/2 \quad -\infty < F < \infty$$

at each receiver. The noises are independent from receiver to receiver. We observe $\mathbf{r}(t) = [r_1(t) \ \dots \ r_M(t)]$, $0 \leq t \leq T$. Then the probability density function is

$$p(\mathbf{r}(t); x_T, y_T) = c_0 \exp \left((-1/N_0)Q \right)$$

where c_0 is a constant and

$$Q = \sum_{i=1}^M \int_0^T (r_i(t) - A_i s(t - \tau_i))^2 dt$$

and where A_i and τ_i both depend on the emitter position. The log-likelihood function is

$$L(x_T, y_T) = (-1/N_0)Q(x_T, y_T)$$

To find the CRLB, the FIM matrix is

$$\mathcal{I}(x_T, y_T) = (1/N_0)E \begin{bmatrix} \frac{\partial^2 Q}{\partial x_T^2} & \frac{\partial^2 Q}{\partial x_T \partial y_T} \\ \frac{\partial^2 Q}{\partial x_T \partial y_T} & \frac{\partial^2 Q}{\partial y_T^2} \end{bmatrix}$$

Now,

$$\frac{\partial Q}{\partial x_T} = -2 \sum_{i=1}^M \int_0^T (r_i(t) - A_i s(t - \tau_i)) \left[-A_i \frac{\partial s(t - \tau_i)}{\partial x_T} - \frac{\partial A_i}{\partial x_T} s(t - \tau_i) \right] dt$$

\Rightarrow

$$\frac{\partial^2 Q}{\partial x_T^2} = -2 \sum_{i=1}^M \int_0^T (r_i(t) - A_i s(t - \tau_i)) \frac{\partial}{\partial x_T} \left[-A_i \frac{\partial s(t - \tau_i)}{\partial x_T} - \frac{\partial A_i}{\partial x_T} s(t - \tau_i) \right] dt +$$

$$-2 \sum_{i=1}^M \int_0^T \left[-A_i \frac{\partial s(t - \tau_i)}{\partial x_T} - \frac{\partial A_i}{\partial x_T} s(t - \tau_i) \right] \left[-A_i \frac{\partial s(t - \tau_i)}{\partial x_T} - \frac{\partial A_i}{\partial x_T} s(t - \tau_i) \right] dt$$

After taking the expected value, the first integral is zero. Thus,

$$-E \left(\frac{\partial^2 Q}{\partial x_T^2} \right) = 2 \sum_{i=1}^M \int_0^T \left[A_i \frac{\partial s(t - \tau_i)}{\partial x_T} + \frac{\partial A_i}{\partial x_T} s(t - \tau_i) \right] \left[A_i \frac{\partial s(t - \tau_i)}{\partial x_T} + \frac{\partial A_i}{\partial x_T} s(t - \tau_i) \right] dt$$

The other terms are found similarly so that

$$\mathcal{I}(x_T, y_T) = (2/N_0) \sum_{i=1}^M \begin{bmatrix} \int_0^T g_{11}^{(i)}(x_T, y_T) dt & \int_0^T g_{12}^{(i)}(x_T, y_T) dt \\ \int_0^T g_{21}^{(i)}(x_T, y_T) dt & \int_0^T g_{22}^{(i)}(x_T, y_T) dt \end{bmatrix}$$

where for example

$$g_{21}^{(i)}(x_T, y_T) = \left[A_i \frac{\partial s(t - \tau_i)}{\partial y_T} + \frac{\partial A_i}{\partial y_T} s(t - \tau_i) \right] \left[A_i \frac{\partial s(t - \tau_i)}{\partial x_T} + \frac{\partial A_i}{\partial x_T} s(t - \tau_i) \right]$$

Converting to the frequency domain by using

$$\int_0^T g(t)h(t) dt = \int_{-\infty}^{\infty} G^*(F)H(F) dF$$

We have upon letting $x_T = \eta_1$ and $y_T = \eta_2$ so that

$$\int_0^T g_{mn}^{(i)} dt = \int_{-\infty}^{\infty} \left(A_i \mathcal{F} \left\{ \frac{\partial s(t - \tau_i)}{\partial \eta_m} \right\} + \frac{\partial A_i}{\partial \eta_m} \mathcal{F} \{ s(t - \tau_i) \} \right)^* \left(A_i \mathcal{F} \left\{ \frac{\partial s(t - \tau_i)}{\partial \eta_n} \right\} + \frac{\partial A_i}{\partial \eta_n} \mathcal{F} \{ s(t - \tau_i) \} \right) dF$$

But $\mathcal{F}\{s(t - \tau_i)\} = S(F) \exp(-j2\pi F \tau_i)$ and $\mathcal{F} \left\{ \frac{\partial s(t - \tau_i)}{\partial \eta_m} \right\} = S(F) \frac{\partial}{\partial \eta_m} \exp(-j2\pi F \tau_i)$

$$\begin{aligned} \int_0^T g_{mn}^{(i)} dt &= \int_{-\infty}^{\infty} \left[A_i S^*(F) \frac{\partial}{\partial \eta_m} \exp(j2\pi F \tau_i) + \frac{\partial A_i}{\partial \eta_m} S^*(F) \exp(j2\pi F \tau_i) \right] \\ &\quad \left[A_i S(F) \frac{\partial}{\partial \eta_n} \exp(-j2\pi F \tau_i) + \frac{\partial A_i}{\partial \eta_n} S(F) \exp(-j2\pi F \tau_i) \right] dF \\ &= \int_{-\infty}^{\infty} |S(F)|^2 \left(A_i \frac{\partial \tau_i}{\partial \eta_m} (j2\pi F) \exp(j2\pi F \tau_i) + \frac{\partial A_i}{\partial \eta_m} \exp(j2\pi F \tau_i) \right) \\ &\quad \left(A_i \frac{\partial \tau_i}{\partial \eta_n} (-j2\pi F) \exp(-j2\pi F \tau_i) + \frac{\partial A_i}{\partial \eta_n} \exp(-j2\pi F \tau_i) \right) dF \\ &= \int_{-\infty}^{\infty} |S(F)|^2 \left(A_i (j2\pi F) \frac{\partial \tau_i}{\partial \eta_m} + \frac{\partial A_i}{\partial \eta_m} \right) \left(A_i (-j2\pi F) \frac{\partial \tau_i}{\partial \eta_n} + \frac{\partial A_i}{\partial \eta_n} \right) dF \\ &= \int_{-\infty}^{\infty} |S(F)|^2 A_i^2 (2\pi F)^2 \frac{\partial \tau_i}{\partial \eta_m} \frac{\partial \tau_i}{\partial \eta_n} dF + \int_{-\infty}^{\infty} |S(F)|^2 \frac{\partial A_i}{\partial \eta_m} \frac{\partial A_i}{\partial \eta_n} dF \end{aligned}$$

The cross terms are zero since

$$\int_{-\infty}^{\infty} F |S(F)|^2 dF = 0$$

So,

$$\int_0^T g_{mn}^{(i)} dt = \int_{-\infty}^{\infty} (2\pi F)^2 |S(F)|^2 dF A_i^2 \frac{\partial \tau_i}{\partial \eta_m} \frac{\partial \tau_i}{\partial \eta_n} + \mathcal{E} \frac{\partial A_i}{\partial \eta_m} \frac{\partial A_i}{\partial \eta_n}$$

where

$$\mathcal{E} = \int_{-\infty}^{\infty} |S(F)|^2 dF$$

is the energy of the signal.

$$\int_0^T g_{mn}^{(i)} dt = \mathcal{E} \bar{F}^2 A_i^2 \frac{\partial \tau_i}{\partial \eta_m} \frac{\partial \tau_i}{\partial \eta_n} + \mathcal{E} \frac{\partial A_i}{\partial \eta_m} \frac{\partial A_i}{\partial \eta_n}$$

where

$$\bar{F}^2 = \frac{\int_{-\infty}^{\infty} (2\pi F)^2 |S(F)|^2 dF}{\int_{-\infty}^{\infty} |S(F)|^2 dF}$$

is the mean-square bandwidth. But

$$\tau_i = (1/c) \sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}$$

\Rightarrow

$$\frac{\partial \tau_i}{\partial x_T} = (1/c) \frac{(x_T - x_i)}{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}} = (1/c) \cos \psi_i$$

and $\frac{\partial \tau_i}{\partial y_T} = (1/c) \sin \psi_i$ where ψ_i is the angle of sensor i from the positive x-axis measured at the emitter as shown in Figure D.1. Hence

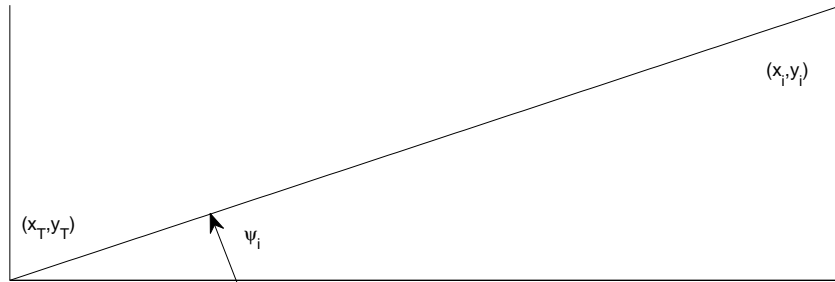


Figure D.1. Definition of the angle ψ_i .

$$[\mathcal{I}(x_T, y_T)]_{mn} = \frac{1}{(N_0/2)} \sum_{i=1}^M \left\{ \mathcal{E} \bar{F}^2 A_i^2 (1/c^2) \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix}_{mn} + \mathcal{E} \frac{\partial A_i}{\partial \eta_m} \frac{\partial A_i}{\partial \eta_n} \right\}$$

or finally we have

$$\begin{aligned} \mathcal{I}(x_T, y_T) &= \frac{\varepsilon}{(N_0/2)} \frac{\bar{F}^2}{c^2} \sum_{i=1}^M A_i^2(x_T, y_T) \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} + \\ &\quad \frac{\varepsilon}{(N_0/2)} \sum_{i=1}^M \begin{bmatrix} \left(\frac{\partial A_i}{\partial x_T} \right)^2 & \left(\frac{\partial A_i}{\partial x_T} \right) \left(\frac{\partial A_i}{\partial y_T} \right) \\ \left(\frac{\partial A_i}{\partial y_T} \right) \left(\frac{\partial A_i}{\partial x_T} \right) & \left(\frac{\partial A_i}{\partial y_T} \right)^2 \end{bmatrix} \end{aligned} \quad (\text{D.1})$$

D.1 With no Obstacles

Assume direct path to all sensors and no multipath. So, the A_i 's, the attenuation factors can be modeled as inversely proportional to the distance. i.e,

$$A_i(x_T, y_T) = \frac{G_i}{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}} = \frac{G_i}{R_i}$$

where G_i is a constant.

$$\Rightarrow \frac{\partial A_i(x_T, y_T)}{\partial x_T} = \frac{-G_i}{R_i^2} \frac{\partial R_i(x_T, y_T)}{\partial x_T} = \frac{-G_i}{R_i^2} \frac{(x_T - x_i)}{\sqrt{(x_T - x_i)^2 + (y_T - y_i)^2}} = \frac{-G_i}{R_i^2} \cos \psi_i$$

and similarly,

$$\frac{\partial A_i(x_T, y_T)}{\partial y_T} = \frac{-G_i}{R_i^2} \sin \psi_i$$

Putting these in (D.1), we have

$$\begin{aligned} \mathcal{I}(x_T, y_T) &= \frac{\varepsilon}{(N_0/2)} \sum_{i=1}^M \left\{ \frac{\bar{F}^2}{c^2} \cdot \frac{G_i^2}{R_i^2} \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} + \right. \\ &\quad \left. \frac{G_i^2}{R_i^4} \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} \right\} \\ &= \frac{\varepsilon}{(N_0/2)} \sum_{i=1}^M \left(\frac{\bar{F}^2}{c^2} \cdot \frac{G_i^2}{R_i^2} + \frac{G_i^2}{R_i^4} \right) \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} \end{aligned}$$

D.2 With an Obstacle

Suppose that there is an obstacle B at (x_B, y_B) . This induces an azimuth modulation say $f_B(x_T, y_T)$. So the attenuation factors can be written as,

$$A_i(x_T, y_T) = \left(\frac{G_i}{R_i} \right) f_B(x_T, y_T)$$

\Rightarrow

$$\begin{aligned} \frac{\partial A_i}{\partial x_T} &= \left(-\frac{G_i}{R_i^2} \right) f_B(x_T, y_T) \frac{\partial R_i(x_T, y_T)}{\partial x_T} + \left(\frac{G_i}{R_i} \right) \frac{\partial f_B(x_T, y_T)}{\partial x_T} \\ &= \left(-\frac{G_i}{R_i^2} \right) f_B(x_T, y_T) \cos \psi_i + \left(\frac{G_i}{R_i} \right) \frac{\partial f_B(x_T, y_T)}{\partial x_T} \end{aligned}$$

And similarly,

$$\frac{\partial A_i}{\partial y_T} = \left(-\frac{G_i}{R_i^2}\right) f_B(x_T, y_T) \sin \psi_i + \left(\frac{G_i}{R_i}\right) \frac{\partial f_B(x_T, y_T)}{\partial y_T}$$

Now,

$$\begin{aligned} \left(\frac{\partial A_i}{\partial x_T}\right)^2 &= \left(-\frac{G_i}{R_i^2}\right)^2 \left[f_B(x_T, y_T) \cos \psi_i - R_i \frac{\partial f_B(x_T, y_T)}{\partial x_T} \right]^2 \\ &= \left(\frac{G_i^2}{R_i^4}\right) f_B^2(x_T, y_T) \left[\cos^2 \psi_i - R_i \frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right]^2 \\ &= \left(\frac{G_i^2}{R_i^4}\right) f_B^2(x_T, y_T) \left[\cos^2 \psi_i + R_i^2 \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right)^2 \right. \\ &\quad \left. - 2R_i \cos \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right) \right] \end{aligned}$$

Similarly, we have

$$\begin{aligned} \left(\frac{\partial A_i}{\partial y_T}\right)^2 &= \left(\frac{G_i^2}{R_i^4}\right) f_B^2(x_T, y_T) \left[\sin^2 \psi_i + R_i^2 \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right)^2 \right. \\ &\quad \left. - 2R_i \sin \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right) \right] \end{aligned}$$

and

$$\begin{aligned} \left(\frac{\partial A_i}{\partial x_T}\right) \left(\frac{\partial A_i}{\partial y_T}\right) &= \left(\frac{G_i^2}{R_i^4}\right) f_B^2(x_T, y_T) \{ \cos \psi_i \sin \psi_i + \\ &\quad R_i^2 \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right) \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right) \\ &\quad - R_i \cos \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right) - R_i \sin \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right) \} \end{aligned}$$

So, we have the FIM as

$$\begin{aligned} \mathcal{I}(x_T, y_T) &= \frac{\varepsilon}{(N_0/2)} \sum_{i=1}^M \left\{ \left(\frac{\bar{F}^2}{c^2} \cdot \frac{G_i^2}{R_i^2} + \frac{G_i^2}{R_i^4} \right) f_B^2(x_T, y_T) \begin{bmatrix} \cos^2 \psi_i & \sin \psi_i \cos \psi_i \\ \sin \psi_i \cos \psi_i & \sin^2 \psi_i \end{bmatrix} \right. \\ &\quad \left. + \left(\frac{G_i^2}{R_i^4} \right) f_B^2(x_T, y_T) \mathbf{J}_i \right\} \end{aligned} \quad (\text{D.2})$$

where the matrix \mathbf{J}_i is given by

$$[\mathbf{J}_i]_{11} = \left[R_i^2 \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right)^2 - 2R_i \cos \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right) \right]$$

$$\begin{aligned}
[\mathbf{J}_i]_{12} &= \left[R_i^2 \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right) \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right) - R_i \cos \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right) \right. \\
&\quad \left. - R_i \sin \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial x_T} \right) \right] \\
&= [\mathbf{J}_i]_{21} \\
[\mathbf{J}_i]_{22} &= \left[R_i^2 \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right)^2 - 2R_i \sin \psi_i \left(\frac{\partial \ln f_B(x_T, y_T)}{\partial y_T} \right) \right] \quad (\text{D.3})
\end{aligned}$$

LIST OF SYMBOLS, ABBREVIATIONS, AND ACRONYMS

AENR	Average energy-to-noise ratio
ASNR	Average signal-to-noise ratio
CAM	Complex cross-ambiguity matrix
CRLB	Cramer-Rao lower bound
FDOA	Frequency difference of arrival
FIM	Fisher information matrix
LPI	Low probability of intercept
LS	Least squares
LSE	Least squared error
MLE	Maximum likelihood estimator
NP	Neyman-Pearson
PDF	Probability density function
RMS	Root mean square
SNR	Signal-to-noise ratio
TDOA	Time difference of arrival
TOA	Time of arrival
UMP	Uniformly most powerful
WLS	Weighted least squares
WLSE	Weighted least squares error
(v_x, v_y, v_z)	The velocity of the target in the x , y and z directions
(x_i, y_i)	Location of sensor i
(x_T, y_T, z_T)	The x , y and z coordinates of the target
$\frac{N_0}{2}$	Noise spectral density
ψ_i	Angle of sensor i from the positive x-axis measured at the emitter

σ_i^2	Noise variance at sensor i
τ_i	Unknown time of arrival of signal at sensor i
\mathbf{A}	$M \times 1$ vector of the unknown attenuation factors
\mathbf{A}'	$M - 1 \times 1$ vector of the unknown relative attenuation factors
\mathbf{B}	Complex cross-ambiguity matrix
\mathbf{I}_N	An $N \times N$ identity matrix
\mathbf{P}	$N \times N$ permutation matrix
\mathbf{W}	$N \times N$ DFT matrix
\tilde{A}	Complex attenuation factor at sensor i
A_i	Unknown attenuation factor at sensor i
a_i, b_i	Fourier coefficients of the signal
c	Propagation speed of signal
d_{\max}	Distance between the farthest pair of sensors
F_0	Fundamental frequency of the Fourier series
F_s	Sampling frequency
k_i	Discrete Doppler shift at sensor i
M	Number of sensors
N	Number of signal samples at each sensor.
n_i	Discrete time delay at sensor i
P_D	Probability of detection
P_{FA}	Probability of false alarm
R_i	Distance from the emitter to the i th sensor
r_i	Signal received at sensor i
s	Transmitted Signal
T	Length of observation interval
t_0	Unknown transmission time

T_s	Non-zero length of the signal
w_i	Additive Gaussian random process at sensor i
\mathcal{E}_{si}	Signal energy at sensor i
\mathcal{I}	Information matrix
\mathcal{I}_θ	Fisher information matrix of the unknown parameter vector θ
η	3×1 vector of the unknown emitter location coordinates and the transmission time
ϕ	$2N - 1 \times 1$ vector of the unknown Fourier coefficients
τ	$M \times 1$ vector of the unknown TOAs
τ'	$M - 1 \times 1$ vector of the unknown TDOAs
$\mathbf{h}(t)$	$2N - 1 \times 1$ vector as defined in Appendix B.3